South American Cacti in Time and Space: Studies on the Diversification of the Tribe Cereeae, with Particular Focus on Subtribe Trichocereinae (Cactaceae)

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde

(Dr.sc.nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Anita Lendel

aus

Kroatien

Promotionskomitee:

Prof. Dr. H. Peter Linder (Vorsitz)

PD. Dr. Reto Nyffeler

Prof. Dr. Elena Conti

Zürich, 2013

Table of Contents

Acknowledgments

I really believe that no one can go through the process of doing a PhD and come out without being changed at a very profound level. Not necessarily because of the PhD itself but because of the surroundings, circumstances, pressure, doubts, questioning and reflections one experiences during that process. The same stands for me. Being away from home for the very first time I learned a lot; about the world but mostly about myself. I've visited beautiful places, met some extraordinary people, widened my personal limits and found the strength I never knew I had in me. And none of this would be possible without many people that deserve to be mentioned here. One person that stands out is Merran Matthews who was a constant perpetuator for me during this process. She found most unorthodox ways to encourage me and "keep me in the game" for which I stay deeply grateful to her. Aelys Humphreys was there for me to discuss within deepest details various thematic overwhelming me at the time – from family values to simplifying approaches I took in a particular project. I profited a lot from conversations with her and thank her for that. Certainly this PhD would never see the light of the day without being supervised by experienced and competent people like Reto Nyffeler, Peter Linder and Urs Eggli. I thank Reto for taking me along this journey and not giving up on me, Peter for supporting this project and me as a person and Urs for generously sharing his knowledge on cacti. Alexandre Antonelli was a "distant player" who would occasionally, still just when needed, bring in an idea or put things into perspective - and that would fuel me to keep on going. Elena Conti was always happy to share her experiences, connected either to science or family issues and I thank her for that. Expert in a floral morphology and with a profound knowledge of angiosperm fossils, Peter Endress was always happy to share his knowledge and together with Merran has provided multiple occasions for me to work on their projects, helping me to stay in Switzerland, financially sustain my thesis and learn valuable techniques, that in many days felt like much needed meditation. Niklaus Müller was always happy to help when my limited knowledge in informatics did not allowed me to continue and I have a feeling that Alex Bernhard was more than amused with my questions on Adobe Illustrator. Dirk Karger is thanked for translating thesis summary to German.

I was lucky enough to discuss methods, techniques and approaches with people that I got to know in the Institute of Systematic Botany: Timo Van der Niet, Jurriaan de Vos, Chloé Galley and Mike Pirie are thanked for their patience while willing to teach and help me with stubborn programs - in many occasions. Marlon Machado was the person who opened up the world of cacti to me. By taking me to meetings of the cactus and succulent societies he made me discover how many people are passionate by this group of plants. It made a great impression on me to see and to meet in person members of the public, non-scientists who actually read our papers and base their knowledge about cacti on our findings – a kind of last stop in a world of science, but unfortunately a stop that many scientists never manage to meet or somehow forget exists. Cyril Guibert was kind enough to travel with me to

California and I thank him for taking the wheel through Los Angeles. I shared in many occasions the load of being a PhD student with Cécile Allanic, Ljiljana Krstin, Lukas Taxböck, Catarina Rydin, Jason Grant and Claire Arnold, and I thank them for listening and encouraging me to continue. Coming to work seemed somehow easier knowing that Gabriele Salvo, Julien Bachelier and Serge Haemmerli will be working next to me and depending on the need will be ready to help, share, or laugh with me. With many of the above mentioned people I shared more than only work - together we partied, laughed and cried. With some of them I stayed in contact, with others less, but regardless I consider them to be my dearest friends, with whom I shared a part of my life and if and whenever our life paths cross again they will never be strangers for me.
And last but not the least the people without whom this PhD would never be completed, and to whom I am indebt the most, are my family. Thank you Pascal for supporting me on this long and painful journey. Thank you Isabelle for being my daughter. My parents, Marica and Mijo together with my brother Dragan Lenđel were more than great in believing in me and giving me opportunities to work by taking care of Isabelle. For this task also many neighbors and friends jumped in and for it I thank Vania Widmer, Jacques Rueger, Chloé Tornay, Andreas Sonderegger, Caroline Barras, and Isabelle Chablais. Family Burkitt is thanked for providing a cosy place to stay at, during the days (nights) before submitting the thesis.

Introduction

Cacti are among the most conspicuous and characteristic plants of arid and semiarid areas in the New World. The family Cactaceae is remarkable for its great diversity of specialized growth form characteristics, floral morphology, and pollination syndromes. The scientific study of Cactaceae, consisting of about 130 genera and 1850 species (Nyffeler and Eggli 2010b) has long been neglected, in spite of many fascinating aspects concerning their evolutionary biology. Even though there are numerous cactus books dedicated to amateurs, the number of detailed taxonomic studies of genera and species groups are very limited (e.g. Leuenberger 1986, Taylor 1991, Zappi 1994, Leuenberger 1997, Nyffeler 1998). The number of names, in most cases not adequately typified, exceeds more than 10 times the number of supposedly „acceptable“ entities and makes the work for monographs very cumbersome (e.g., Hunt 1991). During the past, a wide range of contrasting conceptions took turns, either preferring a more conservative and lumper's approach or arguing for a splitters approach. This fact can be explained by the difficulties to arrive at a sound biological classification system by using morphological characters. Furthermore, a lack of thorough phylogenetic studies prevented so far to resolve controversial opinions on relationships and classification. Overall, a small number of molecular phylogenetic studies have been devoted to Cactaceae up to now, resolving phylogenetic relationships between cacti and its closest relatives (Nyffeler 2007, Nyffeler and Eggli 2010a), among major clades within the cactus family (Nyffeler 2002, Ocampo and Columbus 2010, Bárcenas, et al. 2011, Hernández-Hernández, et al. 2011), at the base of the cactus phylogeny (Butterworth and Wallace 2005, Edwards, et al. 2005, Butterworth and Edwards 2008), as well as among various suprageneric and generic groups of Cactoideae and Opuntioideae (e.g., Demaio, et al. 2011, Mosti, et al. 2011, Schlumpberger and Renner 2012).

The typical cactus growth form, characterized by the absence of leaves, stems that stay green and photosynthetically active for several seasons, spine clusters associated with side branches, and a reduced branching pattern is also found in other, unrelated groups of eudicot plant families (i.e., Apocynaceae, Euphorbiaceae). It represents one of the classical textbook example for "convergent evolution" for adaption to a prominent ecological factor (i.e., temporal aridity; Niklas 1997, Futuyma 1998). The traditional idea put forward by Buxbaum (1956) in his „law of the abbreviation of the vegetative phase" is that there is a general „trend" in cactus evolution leading from branched, arborescent-columnar forms to unbranched, solitary, globular types (e.g. Buchheim 1964, Cullman, et al. 1984, see Nyffeler and Eggli 2010b). This idea concerning cactus evolution (Fig.1) and relationships is still widely prevailing and is reflected in current classification systems (e.g. Buchmann 1964, Cullmann, et al. 1984) .

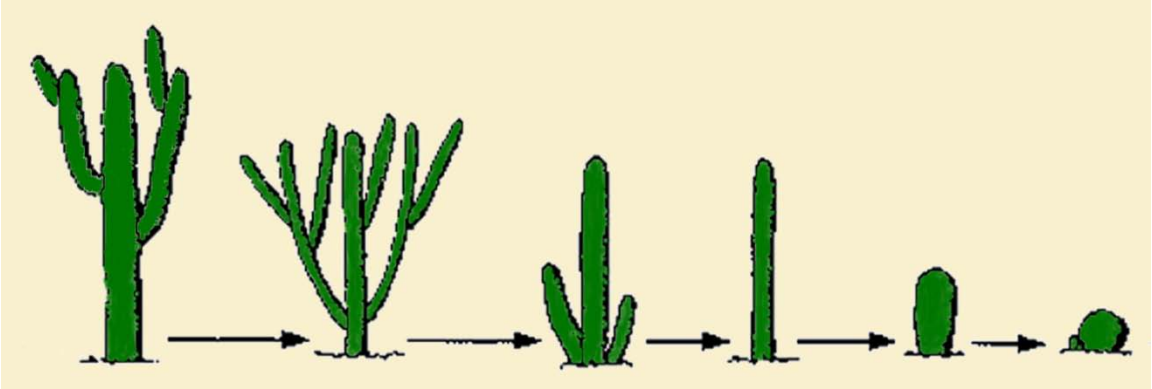


Figure 1. Schema of the prevailing concept on cactus evolution, showing reduction of the vegetative phase.

However, recent molecular systematic analyses (e.g., Fig. 2; Hernández-Hernández, et al. 2011, Schlumpberger and Renner 2012) documented evidence that these ideas concerning the pattern of growth form evolution in cacti need to be revised (see also Griffith 2004a, Griffith 2004b).

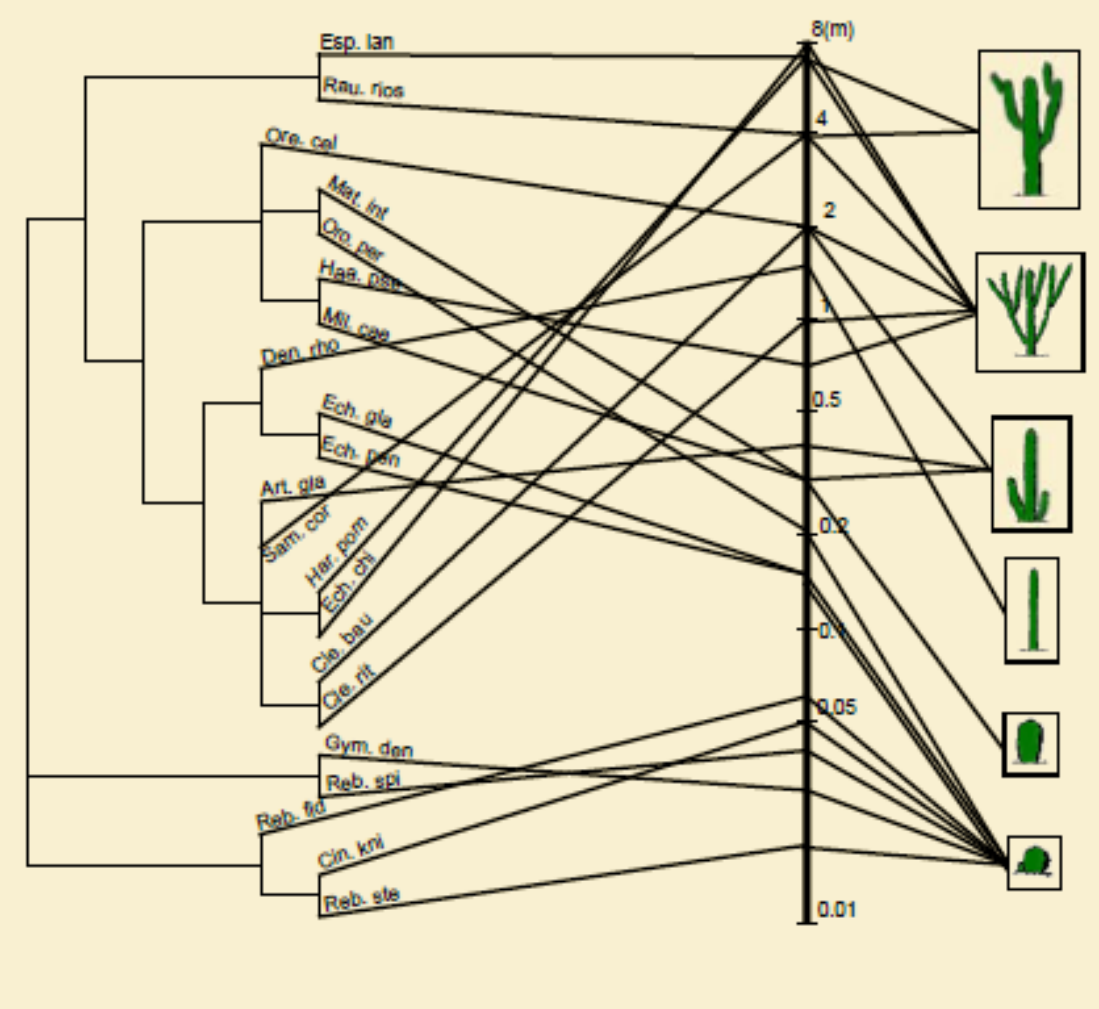


Figure 2. Summary figure illustrating the large amount of homoplasy in the evolution of growth form characteristics in species of the tribe Trichocereeae sensu Hunt et al. (2006).

This preliminary analysis at the outset of the dissertation project was based on phylogenetic analysis

of 80 species of cacti performed using maximum parsimony method on the data set combined of four chloroplast markers (*trnK*/*matK*, *rpl16*, *rps16* and *trnSG*). Only part of the tree, the clades of interest of the consensus tree (out of 504 most parsimonious trees) are shown. For each taxon on a phylogenetic tree an approximate growth height is noted and its log value and habitat type is associated with it. Esp. lan = *Espostoa lanata*; Rau. rios = *Rauhocereus riosaniensis*; Ore. cel = *Oreocereus celsianus*; Mat. int = *Matucana intertexta*; Oro. per = *Oroya peruviana*; Haa. pse = *Haageocereus pseudomelanostele*; Mil. cae = *Mila caespitosa*; Den. rho = *Denmoza rhodacantha*; Ech. gla = *Acanthocalycium thionanthum*; Ech. pen = *Lobivia pentlandii*; Art.gla = *Arthrocereus glaziovii*; Sam.cor = *Samaipaticereus corroanus*; Har. pom = *Harrisia pomanensis*; Ech.chi = *Echinopsis chiloensis*; Cle.bau = *Cleistocactus baumanii*; Cle. rit = *Cleistocactus ritteri*; Gym. den = *Gymnocalycium denudatum*; Reb. spi = *Aylostera_fiebrigii*; Reb. fid = *Weingartia fidana*; Cin. kni = *Weingartia cintia*; Reb.ste = *Weingartia steinbachii*.

Furthermore, recent studies on the patterns of floral evolution and pollination syndromes document the large number of evolutionary shifts between flowers adapted to pollination by bats, bees, hawkmoths, or hummingbirds (e.g., Schlumpberger and Renner 2012). Since floral characters have provided important base for generic and suprageneric classification in the past, hence, such classification systems do not reflect phylogenetic relationships.

The origin of Cactaceae has puzzled researchers and the broader public for centuries (e.g., Hunt and Taylor 1990, Hershkovitz and Zimmer 1997) and has led to speculations about their age of origin. Traditionally, rather old ages of 110–90 Ma (e.g., Gibson and Nobel 1986, Mauseth 1990, Butterworth and Edwards 2008: "deep in the Cretaceous") have been invoked for the family in order to allow for the evolution of their distinctive features and geographical distribution. Early molecular phylogenetic studies (e.g. Hershkovitz and Zimmer 1997) refer the origin of the cacti to the mid-Tertiary, c. 30 mya on the basis of sequence divergence of the investigated marker and expected substitution rate as inferred from other angiosperms. Although more recent studies (Ocampo and Columbus 2010, Arakaki, et al. 2011), focusing on the phylogenies of the cacti and their relatives, discuss the possible age and the place of origin of these groups, studies inferring a more detailed temporal or spatial scenario of the evolutionary history of the various lineages of cacti, and also considering the substantial methodological and analytical challenges for this group, have not been conducted so far. The lack of a fossil record for this group of plants as well as their close relatives is certainly one of the many intricacies while studying this group. Fossils provide direct evidence for the age of the lineage to which they belong. However, fossilization in arid environment is hampered by the lack of water containing mineral solvents that are required for the fossilization of plant parts.

The present thesis addresses, on the basis of comparative molecular phylogenetic analyses, the relationships among the major subclades and genera of the prominent “BCT” clade of Cactoideae (Nyffeler 2002), i.e. Cereeae s.l. sensu Nyffeler & Eggli (2010b), with a particular focus on the subtribe Trichocereinae (Buxbaum 1958, Nyffeler and Eggli 2010b). The tribal and subtribal classification is revised and the pattern of geographical diversification is briefly discussed. Based on this phylogenetic study, the evolutionary

scenarios of several floral characters and floral syndromes within the tribe Cereeae s.l. are explored and ancestral character reconstruction is conducted. Furthermore, a prominent part of this thesis is devoted to the estimation of divergence times of different lineages of cacti and to it associated methodological and analytical challenges.

Molecular systematic studies of the family Cactaceae indicate that a large part of the globular and columnar cacti of South America constitute a well-supported monophyletic group that is widely known as “BCT” clade (Nyffeler 2002, Wallace and Gibson 2002, Crozier 2005, Ritz, et al. 2007, Bárcenas, et al. 2011, Hernández-Hernández, et al. 2011). This distinct clade, the largest and morphologically as well as ecologically most diverse group of South American Cactoideae, consists of some 40 genera and almost 600 species, which previously have been classified in the three separate tribes Browningieae, Cereeae, and Trichocereeae (e.g. Endler and Buxbaum 1973, Barthlott and Hunt 1993, Anderson 2001, Anderson 2005). The tribes of the BCT clade are defined primarily based on morphological characters of flowers and fruits, as well as biogeographical considerations (e.g. Buxbaum 1969, Barthlott and Hunt 1993, Anderson 2001, Anderson 2005, Hunt, et al. 2006) and a close relationship between the members of the tribes has been proposed for a long time (e.g. Buxbaum 1969, Barthlott 1979, Barthlott 1988, Barthlott and Hunt 1993). The traditional tribe Trichocereeae comprises about 25 genera and some 300 species (Table 1), including some of the most diverse and attractive cacti from southern South America (i.e., *Cleistocactus*, *Echinopsis*, *Rebutia*). Species of Trichocereeae are found at localities from sea level to more than 4000 m in various different habitats ranging from arid areas in the Atacama desert of Peru to savannah areas in Uruguay and southern Brazil. With a wide spectrum of growth forms ranging from trees and shrubs to caespitose or single globular stems and a remarkable floral diversity often correlated to pollination syndrome, this clade is a typical example for the misleading information for relationships provided by characters such as growth form and floral characters (e.g., *Echinopsis* s.s., *Lobivia*, *Pseudolobivia* in Backeberg 1966). Previous molecular phylogenetic investigations (e.g., Nyffeler 2002) showed the traditional tribe Trichocereeae to be a prominent part of the “BCT” clade, however, outlined a need for a more detailed investigation based on an expanded taxonomic and molecular marker sampling.

The focus of the first part of this thesis (Chapter 1) lies on establishing a hypothesis on phylogenetic relationships between well supported monophyletic taxa and discerning a well supported generic and suprageneric classification of the “BCT” clade. In particular, by testing for the support of the alternative phylogenetic hypothesis we investigate the phylogenetic relationships and taxonomic state of a number of its genera and suprageneric clades to arrive at an informed decision for or against formally recognizing a given clade in a revised classification system. In the light of the phylogenetic hypothesis a comparative investigation on the diversification patterns of five morphological characters (i.e., perianth segment color, floral symmetry, position of flowers along the stem as well as presence/absence of a cephalium, presence of hairs and bristles of the indumentum of the pericarpel and perianth tube and the time of anthesis) traditionally used for classification into tribes and genera as well as floral syndrome are studied and ancestral character state

reconstruction provides insights on the amount of parallel or convergent evolution in these characteristics (Chapter 2).

In order to better understand how the diversity of such a prominent group of cacti came along, it is important to elucidate the timeframe in which cacti originated, as well as the age of divergence of its major lineages. Due to the lack of a fossil record for cacti and their closest relatives, age estimation studies need to expand the study group to include taxa with a fossil record for calibration. These research questions are further investigated in two different studies. On the one hand, a broad-scale investigation (Chapter 3) on the diversification of succulent plant lineages, the origin of cacti based on age estimation is linked to particular climatological and geological events in North and South America. The time of origin of extant cacti and the age of several prominent subclades is inferred in this study in a Bayesian framework in a two-step approach, employing a secondary calibration approach. Furthermore, rates of diversification for several clades of cacti are estimated to investigate the possible link between global changes in environmental conditions with the onset of prominent radiations in separate lineages adapted to arid conditions. In a different study (Chapter 4) we explore the estimation of divergence times in the particularly challenging case if the taxonomic group of interest has a scarce or missing fossil record. The "molecular dating paradox" refers to the situation, well exemplified by the cacti, where taxonomic groups that lack a reliable fossil record and, hence, whose divergence ages would therefore potentially benefit most from the molecular dating are at the same time most sensitive to biases that cannot be adjusted by evidence from fossils. The performances of three different approaches to circumvent this problem is evaluated: 1) a "Distant Fossil Calibration approach", where the molecular data set for the study group is expanded to include relatives with a large record of informative fossils; 2) a "Study Group Placeholder approach", whereby age estimates for the study group are derived from a higher level dating analysis that includes only few representatives of the study group as placeholders; and 3) a "Secondary Calibration approach", which uses for time calibration estimated age derived from a previous dating analysis, rather than fossils. Furthermore, the effect of using different taxon sampling densities and analytical methods is explored. The consecutive analytical steps employed in this study are schematically presented as a flowchart (Fig. 3). Components (i.e., topology, fossils and sequence matrix), resulting from different lines of evidence (e.g., molecular phylogenetics and paleontology) are used as building blocks in establishing an analytical protocol towards estimating the age of cacti. Those analytical steps rely on applying different software packages (i.e. PAUP, r8s, PATHd8 or BEAST; Swofford 2002, Sanderson 2003, Sanderson 2006, Britton, et al. 2007, Drummond and Rambaut 2007) that implement different analytical methods (i.e. ML, NPRS, PL, PATHd8 and UCLN; Sanderson 1997, Sanderson 2002, Swofford 2002, Drummond, et al. 2006, Britton, et al. 2007). This investigation is based on sequence information from a single marker, *matK*, (422 sequences from GenBank, 38 newly generated sequences). Phylogenetic analyses, applying a parsimony approach (using PAUP) as well as a model-based approach in a Bayesian framework (using MrBayes; Huelsenbeck and Ronquist 2001, Ronquist and Huelsenbeck 2003) are preformed in order to yield tree topologies for age

estimation and to compare the results with those derived from a UCLN relaxed clock model implemented in BEAST. The wide range of age estimates derived from 18 different data sets and different analytical methods are compared resulting in a well founded estimate of divergence age for the cacti.

The thesis is finished with a summary and an overview of the insights from the current investigations on this prominent lineage of South American cacti. This concerns in particular more detailed biogeographical investigations, which will greatly profit from the well founded phylogenetic investigations as well as from the age estimation studies.

Literature Cited:

Anderson EF. 2001. The Cactus family. Portland, Oregon (U.S.A.), Timber press, Inc.

Anderson EF. 2005. Das grosse Kakteen-Lexikon. Dörr + Schiller, Stuttgart, Printer Trento S.r.l.,.

Arakaki M, Christin PA, Nyffeler R, Lendel A, Eggli U, Ogburn RM, Spriggs E, Moore MJ, Edwards EJ. 2011. Contemporaneous and recent radiations of the world's major succulent plant lineages. Proc. Natl. Acad. Sci. U. S. A., 108:8379-8384.

Backeberg C. 1966. Das Kakteenlexikon. Enumeration Diagnostica Cactacearum. Jena (DDR), VEB Gustav Fischer Verlag.

Bárcenas RT, Yesson C, Hawkins JA. 2011. Molecular systematics of the Cactaceae. Cladistics, 27:470-489.

Barthlott W. 1979. Cacti. Cheltenham, Stanley Thornes.

Barthlott W. 1988. über die Systematischen Gliederung der Cactaceae. Beitr. Biol. Pflanzen., 63 (1/2):17-40.

Barthlott W, Hunt DR. 1993. The families and genera of vascular plants. In: Kubitzki, K & al. (eds.). Berlin (D) etc., Springer-Verlag, p. 161-197, ills.

Britton T, Anderson CL, Jacquet D, Lundqvist S, Bremer K. 2007. Estimating divergence times in large phylogenetic trees. Syst. Biol., 56:741-752.

Buchheim G. 1964. Cactales. In: Melchior H editor. A. Engler's Syllabus der Pflanzenfamilien. Berlin, Gebrüder Borntraeger, ed. 12, 2:, p. 102-108.

Buchmann G. 1964. Cactaceae. In: Melchior H editor. Engler’s Syllabus der Pflanzenfamilien mit besonderer Berücksichtigung der Nutzpflanzen nebst einer Übersicht über die Florenreiche und Florengebiete der Erde. Berlin, Gebrüder Borntraeger, p. 102-108.

Butterworth C, Edwards EJ. 2008. Investigating *Pereskia* and the Earliest Divergences In Cactaceae. Haseltonia:46-53.

Butterworth CA, Wallace RS. 2005. Molecular phylogenetics of the leafy cactus genus *Pereskia* (Cactaceae). Syst. Bot., 30:800-808.

Buxbaum F. 1956. Das Gesetz der Verkürzung der vegetativen Phase in der Familie der Cactaceae. Österr. Bot. Z., 103:353-362.

Buxbaum F. 1958. The phylogenetic division of the subfamily *Cereoideae*, Cactaceae. Madroño, 14:177-206.

Buxbaum F. 1969. Die Entwicklungswege der Kakteen in Sudamerika. In: Fittkau EJ, Illies J, Klinge H, Schwabe GH, Sioli H editors. Biogeography and ecology in South America. The Hague, The Netherlands, Junk, W. , p. 583–623.

Crozier S. 2005. Systematics of Cactaceae Juss. [electronic resource]: phylogeny, cpDNA evolution, and classification, with emphasis on the genus *Mammillaria* Haw., Doctoral dissertation, The University of Texas at Austin.
Cullman W, Gotz E, Groner G. 1984. Encyclopedia of cacti. Portland, Ore., Timber Press.
Cullmann W, Götz E, Gröner G. 1984. Kakteen: Kultur, Vermehrung und Pflege. Lexikon der Gattungen und Arten. Stuttgart, Ulmer, E.
Demaio PH, Barfuss MHJ, Kiesling R, Till W, Chiapella JO. 2011. Molecular phylogeny of *Gymnocalycium* (Cactaceae): assessment of alternative infrageneric systems, a new subgenus, and trends in the evolution of the genus. American Journal of Botany, 98:1841-1854.
Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. PLoS. Biol., 4:699-710.
Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol., 7.
Edwards EJ, Nyffeler R, Donoghue MJ. 2005. Basal cactus phylogeny: implications of *Pereskia* (Cactaceae) paraphyly for the transition to the cactus life form. American Journal of Botany, 92:1177-1188.
Endler J, Buxbaum F. 1973. Die Pflanzenfamilie der Kakteen. Ein systematischer Wegweiser für Liebhaber und Erwerbszüchter mit einer kompletten Liste der Gattungssynonyme. 3 ed. Minden (D), Albrecht Philler Verlag.
Futuyma DJ. 1998. Evolutionary Biology, 3rd ed. Sinauer Associates, Sunderland, MA.
Gibson AC, Nobel PS. 1986. The cactus primer. Cambridge (US:MA)/London (GB), Harvard University Press.
Griffith MP. 2004a. Early cactus evolution: The postmodern view. Haseltonia:3-11.
Griffith MP. 2004b. What did the first cactus look like? An attempt to reconcile the morphological and molecular evidence. Taxon, 53:493-499.
Hernández-Hernández T, Hernández HM, De-Nova JA, Puente R, Eguiarte LE, Magallon S. 2011. Phylogenetic relationships and evolution of growth form in Cactaceae (Caryophyllales, eudicotyledoneae). American Journal of Botany, 98:44-61.
Hershkovitz MA, Zimmer EA. 1997. On the evolutionary origins of the cacti. Taxon, 46:217-231.
Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics, 17:754-755.
Hunt DR. 1991. Stabilization of names in succulent plants. In: Hawksworth DL editor. Improving the stability of names: needs and options. Königstein, D., Koeltz Scientfic Books, p. 151-155.
Hunt DR, Taylor N, Graham C. 2006. The New Cactus Lexicon. Dorchester (UK), dh books.
Hunt DR, Taylor NP. 1990. The genera of the Cactaceae: progress towards consensus. Bradleya, 8:85-107.
Leuenberger BE. 1986. *Pereskia* (Cactaceae). Mem. New York Bot. Gard., 41:1-141.
Leuenberger BE. 1997. *Maihuenia* - monograph of a Patagonian genus of Cactaceae. Bot. Jahrb. Syst., 119:1-92.
Mauseth JD. 1990. Continental drift, climate and the evolution of the cacti. Cactus and succulent journal (U.S.A.), 62:302-308.
Mosti S, Bandara NL, Papini A. 2011. Further insights and new combinations in *Aylostera* (Cactaceae) based on molecular and morphological data. Pak. J. Bot., 43:2769-2785.
Niklas K. 1997. The Evolutionary Biology of Plants., University of Chicago Press.
Nyffeler R. 1998. The genus *Uebelmannia* (Cactaceae). Bot. Jahrb. Pflanzensystematik, 120:145-163.

Nyffeler R. 2002. Phylogenetic relationships in the cactus family (Cactaceae) based on evidence from trnK/matK and trnL-trnF sequences. American Journal of Botany, 89:312-326.
Nyffeler R. 2007. The closest relatives of cacti: Insights from phylogenetic analyses of chloroplast and mitochondrial sequences with special emphasis on relationships in the tribe Anacampseroteae. American Journal of Botany, 94:89-101.
Nyffeler R, Eggli U. 2010a. Disintegrating Portulacaceae: A new familial classification of the suborder Portulacineae (Caryophyllales) based on molecular and morphological data. Taxon, 59:227-240.
Nyffeler R, Eggli U. 2010b. A farewell to dated ideas and concepts: molecular phylogenetics and a revised suprageneric classification of the family Cactaceae. Schumannia, 6:109-149.
Ocampo G, Columbus JT. 2010. Molecular phylogenetics of suborder Cactineae (Caryophyllales), including insights into photosynthetic diversification and historical biogeography. American Journal of Botany, 97:1827-1847.
Ritz CM, Martins L, Mecklenburg R, Goremykin V, Hellwig Frank H. 2007. The molecular phylogeny of *Rebutia* (Cactaceae) and its allies demonstrates the influence of paleogeography on the evolution of South American montain cacti. American Journal of Botany, 94:1321-1332.
Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics, 19:1572-1574.
Sanderson MJ. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. Mol. Biol. Evol., 14:1218-1231.
Sanderson MJ. 2002. Estimating absolute rates of molecular evolution and divergence times: A penalized likelihood approach. Mol. Biol. Evol., 19:101-109.
Sanderson MJ. 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics, 19:301-302.
Sanderson MJ. 2006. r8s version 1.71. Analysis of rates ("r8s") of evolution., Section of Evolution and Ecology, University of California, Davis.
Schlumpberger BO, Renner SS. 2012. Molecular phylogenetics of *Echinopsis* (Cactaceae): polyphyly at all levels and convergent evolution of pollination modes and growth forms. American Journal of Botany, 99:1335-1349.
Swofford DL. 2002. PAUP*: Phylogenetic Analysis Using Parsimony (*and other methods). Sunderland, MA, Sinauer Associates.
Taylor NP. 1991. The genus *Melocactus* in Central and South America. Bradleya, 9:1-80.
Wallace RS, Gibson AC. 2002. Evolution and Systematics. In: Nobel PS editor. Cacti Biology and Uses, University of California press, p. 1-21.
Zappi D. 1994. *Pilosocereus* (Cactaceae). The genus in Brazil. Succulent Pl. Res., 3:1-160.

Figures and Tables:

Figure 1. Framework of the prevailing idea on cactus evolution directionality: from a treelike or shrubby growth form, through a uni-directional reduction to a more derived dwarf growth form.

Figure 2. Summary figure illustrating the large amount of homoplasy in the evolution of growth form characteristics in species of the tribe Trichocereeae sensu Hunt et al. (2006).

Figure 3. The schematic workflow we employ in this study presented as a flowchart. We employ components (i.e. topology, fossils and matrix), resulting from different fields of science (e.g. molecular phylogenetics and paleontology) as building blocks in constructing our path towards estimating the age of cacti. Those components are analyzed with software (i.e. PAUP, r8s, PATHd8 or BEAST) implementing various methods (i.e. ML, NPRS, PL, PATHd8 and UCLN) and the effects of the relaxed clocks on the age estimates are observed. Additionally, a heuristic search (in PAUP) and Bayesian phylogenetic analyses (in MrBayes) are performed in order to compare the observed topologies with the designed starting topology. Steps shown in gray outline either steps we did not perform in this study (e.g. confidence intervals for NPRS, PL and PATHd8 divergence age estimates), or additional possibilities (e.g. topology comparisons) and alternative ways in estimating divergence ages, that could provide supplementary information.

Table 1. Comparison of classifications for the genera and species of the Trichocereeae / Trichocereinae. Gray box presents accepted genera according to the classification outlined for each of the columns. Genera associated with Trichocereeae / Trichocereinae are listed and the number of accepted (and provisionary accepted) species is outlined according to the associated classification. When the genus was not accepted for a give classification, its synonyms are outlined. Finally a total number of genera and species is counted for each of the classification.

Figure 3

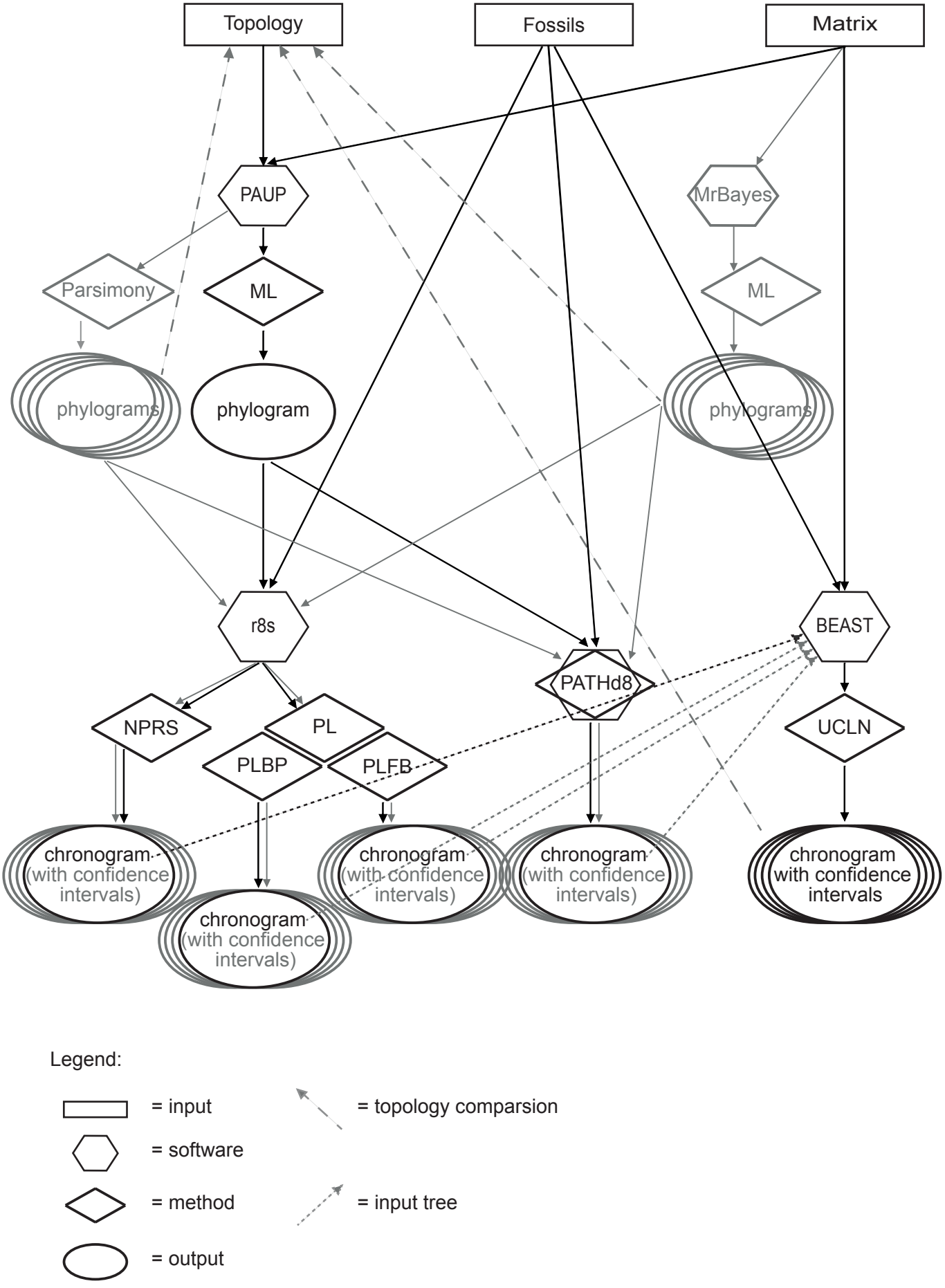

Topology
Fossils
Matrix
PAUP
MrBayes
Parsimony
ML
ML
phylograms
phylogram
phylograms
r8s
BEAST
PATHd8
NPRS
PL
PLBP
PLFB
UCLN
chronogram (with confidence intervals)
chronogram (with confidence intervals)
chronogram (with confidence intervals)
chronogram with confidence intervals
chronogram (with confidence intervals)
Legend:
= input
= software
= method
= output
= topology comparsion
= input tree

Table 1

genera accepted as traditional Trichocereeae / Trichocereinae							number of **species** recognized as traditional Trichocereeae / Trichocereinae				
1958	1973	1993	2001	2005	2006	2010	**genus**	Hunt 1999	Anderson 2001	Anderson 2005	Hunt 2006	Nyffeler & Eggli 2010
							Acanthocalycium	1 (2)	3	5	≡ *Echinopsis, Eriosyce*	5
							Arequipa		≡ *Oreocereus* , *Matucana*	≡ *Oreocereus* , *Matucana* , *Cleistocana* ,	≡ *Oreocereus, Matucana*	≡ *Oreocereus*
							Arthrocereus	4	4	4	4	4
							Borzicactus		≡ *Cleistocactus, Haageocereus, Matucana, Oreocereus, Corryocactus*			ca. 20
							Brachycereus	1	1	-	1	Corryocactinae
							Cephalocleistocactus	(1)	1	1	≡ *Cleistocactus*	1
							Chamaecereus		≡ *Echinopsis*			≡ *Echinopsis*
							Cleistocactus	33 (16)	48 (+1 hybrid)	53 (+1 hybrid)	38	ca. 30 incl. *Akersia* , *Bolivicereus* , *Borzicactella* , *Hildewintera* , *Winterocereus*
							Denmoza	1	1	1	1	1
							Digitorebutia		≡ Rebutia			≡ Rebutia
							Discocactus	6 (1)	7	7	11	Cereinae
							Echinopsis	61 (69)	128 (+1 hybrid)	123 (+2 hybrids)	77	126 incl. *Helianthocereus, Pseudolobivia, Soehrensia, Trichocereus*
							Espostoa	9 (7)	16	17	11	17 incl. *Thrixanthocereus*
							Espostoopsis	1	1	1	excluded from Trichocereeae	1
							Facheiroa	3	3	3	3	Cereinae
							Gymnocalycium	42 (28)	71	78 (+2 hybrids)	49	Rebutiinae
							Haageocereus	13 (8)	20	19	9	12 incl. *Maritimocereus*
							xHaagespostoa	(2)	2	2	not treated	-
							Harrisia	14 (6)	20	21	9	20 *Eriocereus, Estevesia* ?
							Lasiocereus	2	2	2	2	Rebutiinae
							Leocereus	1	1	1	1	Cereinae
							Lobivia		≡ *Echinopsis, Rebutia*			≡ *Echinopsis*
							Loxanthocereus		≡ *Loxanthocereus, Matucana*			≡ *Haageocereus?*
							Matucana	15 (2)	17	17	14	17 incl. *Eomatucana* , *Submatucana*
							Micranthocereus	-	-	-	-	Cereinae
							Mila	3 (1)	1	1	1	1
							Morawetzia			≡ *Oreocereus*		≡ *Oreocereus*
							Oreocereus	5 (4)	9	9	6	9
							Oroya	1 (1)	2	2	2	2
							Pygmaeocereus	2 (1)	3	3	2	≡ *Haageocereus*
							Rauhocereus	1	1	1	1	1
							Rebutia	24 (17)	41	31	29	Rebutiinae
							Samaipaticereus	1	1	1	1	1
							Setiocereus		≡ *Cleistocactus, Browningia*			≡ *Haageocereus?*
							Setiechinopsis		≡ *Echinopsis*			≡ *Echinopsis*
							Soehrensia		≡ *Echinopsis*			≡ *Echinopsis*
							Sulcorebutia		≡ *Rebutia, Neowerdermannia*	16	≡ *Rebutia, Neowerdermannia*	≡ *Weingartia*
							Vatricania					1
							Trichocereus		≡ *Echinopsis, Weberbauerocereus*	≡ *Arthrocereus, Echinopsis, Weberbauerocereus*	≡ *Echinopsis, Weberbauerocereus*	≡ *Echinopsis*
							Weberbauerocereus	7 (1)	8	8	7	8
							Weingartia		≡ *Rebutia, Neowerdermannia*	4	≡ *Rebutia, Neowerdermannia*	Rebutiinae
							Zehntnerella			≡ *Facheiroa*		≡ *Facheiroa*
							Yungasocereus	1	1	1	1	1
22	25	18	27	27	23	20	Total Trichocereae	252 (provisionary accepted=419)	413	430	280	278

1958 = Buxbaum F. 1958. The phylogenetic division of the subfamily cereoideae, cactaceae. Madroño, 14:177-206.
1973 = Endler J, Buxbaum F. 1973. Die pflanzenfamilie der kakteen. Ein systematischer wegweiser für liebhaber und erwerbszüchter mit einer kompletten liste der gattungssynonyme. 3 ed. Minden (D), Albrecht Philler Verlag.
1993 = Barthlott W, Hunt DR. 1993. The families and genera of vascular plants. In: Kubitzki, k & al. (eds.). Berlin (D) etc., Springer-Verlag, p. 161-197, ills.

Hunt DR. 1999. Cites cactaceae checklist. 2 ed. Richmond (GB): Royal Botanic Gardens Kew/Milbourne Port (GB).
Hunt DR, Taylor N, Graham C. 2006. The new cactus lexicon. Dorchester (UK), dh books.
Nyffeler R, Eggli U. 2010. A farewell to dated ideas and concepts: Molecular phylogenetics and a revised suprageneric classification of the family cactaceae. Schumannia, 6:109-149.

Chapter 1.

Phylogenetics and taxonomy of the tribe Cereeae s.l., with particular focus on the subtribe Trichocereinae (Cactaceae – Cactoideae)

Anita Lendel[1], Urs Eggli[2], Reto Nyffeler[*1]

[1] Institut für Systematische Botanik, Zollikerstrasse 107, CH-8008, Zürich, Switzerland
[2] Sukkulenten-Sammlung Zürich, Mythenquai 88, CH-8002 Zürich, Switzerland

Corresponding author. E-mail address: reto.nyffeler@systbot.uzh.ch

Manuscript for Taxon

Abstract

On the basis of four combined cpDNA markers and applying maximum likelihood and maximum parsimony methods, a phylogenetic hypothesis about the relationships among all relevant representatives of the tribes Cereeae and Trichocereeae is presented. These investigations confirm that the two tribes should be combined to form an expanded tribe Cereeae that is classified into three subtribes and three distinct lineages classified as genera *Aylostera* (and possibly *Mediolobivia* as separate genus), *Gymnocalycium*, and *Uebelmannia* of largely unresolved relationships. Particular attention is devoted to a discussion of the implications of this approach to classification in relation to traditional, morphology-based classification. The widely circumscribed genera *Cleistocactus*, *Echinopsis*, and *Rebutia*, are identified as polyphyletic groups that need to be split into monophyletic entities. We use parsimony- and likelihood-based paired-sites tests to investigate whether the molecular data is in conflict to traditional taxonomic concepts. Hence, we propose (1) to recognize *Bolivicereus*, *Borzicactus*, and *Loxanthocereus* in addition to *Cleistocactus*, (2) to recognize *Acanthocalycium*, *Lobivia*, *Setiechinopsis*, and *Trichocereus* in addition to *Echinopsis*, (3) to recognize *Vatricania* in addition to *Espostoa*, and (4) to recognize *Aylostera* and *Weingartia* (possibly also *Mediolobivia* and *Sulcorebuta*) in addition to *Rebutia*. The tribe Cereeae s.l. is basically a South American taxon, with only few species extending to the Caribbean. Subtribe Cereinae is centered in the arid regions of NE Brazil, while the subtribes Rebutiinae and Trichocereinae are prominent on the Altiplano and the foothills of Argentina, Bolivia, Chile, and Peru, with several East/West Andean migrations.

Keywords Cactaceae; Cereeae; classification; molecular phylogeny; South America; test for alternative topologies tests; Trichocereeae

Introduction

The generic and suprageneric classification of cacti is renowned for its intricate history (e.g., Metzing & Kiesling, 2008; Bárcenas et al., 2011). Several contrasting trends and opinions, either preferring a more conservative and lumper's approach or arguing for a splitter's approach, have taken turns during the past century. Part of this wide range of conceptions is explained by the difficulties to discern natural groups, or, in today's view, phylogenetic relationships among different lineages. Particularly, morphological characters proved to be difficult to interpret in order to arrive at well founded classification systems (e.g., Wallace, 1995). Prior to the 1980ies, the widely used horticultural classification systems of Curt Backeberg (Backeberg, 1958-1962, 1966) were in general use, but increasingly recognized as vastly artificial and thus inappropriate to depict phylogenetic relationships. But also the proclaimed phylogenetic classification systems of Franz Buxbaum (e.g., Buxbaum, 1958; Endler & Buxbaum, 1974) were increasingly criticized (e.g., Barthlott, 1979; Gibson & Nobel, 1986).

Later, a working group derived from members of the International Organization for Succulent Plant Study applied a conservative "consensus classification" approach in order to overcome the widely opposed classification systems (Hunt & Taylor; 1986; Hunt & Taylor, 1990; Hunt & al., 2006). In recent years, molecular phylogenetic studies hugely improved our knowledge on supposed phylogenetic relationships for the family Cactaceae (e.g., Nyffeler, 2002; Edwards & al., 2005; Bárcenas et al., 2011; Hérnandez-Hérnandez et al., 2011). Phylogenies have repeatedly been used to derive statements on monophyly or non-monophyly (i.e., paraphyly or polyphyly) of certain genera and suprageneric entities and to call for adjustments (e.g., Nyffeler, 2002; Ritz & al., 2007; Schlumpberger & Renner, 2012). Notwithstanding, only rarely have rigorous tests for the support of alternative phylogenetic hypotheses been applied to arrive at an informed decision for or against a given clade to be recognized in a revised classification system.

All molecular systematic studies of the family Cactaceae indicate that a large part of the globular and columnar cacti of South America constitute a well-supported monophyletic group that is widely known as "BCT clade" (Nyffeler, 2002; Wallace & Gibson, 2002; Croizier, 2005; Ritz & al., 2007; Bárcenas & al., 2011; Hernández-Hernández & al., 2011). This clade consists of some 40 genera and almost 600 species, which have previously been classified in the three tribes Browningieae, Cereeae, and Trichocereeae (e.g., Endler & Buxbaum, 1974; Barthlott & Hunt, 1993; Anderson, 2001, 2005). Recently, Hunt & al. (2006), based on insights from molecular phylogenetic studies, proposed a rearrangement of the genera of the tribe Browningieae to be either included in Cereeae (i.e., *Browningia*, *Stetsonia*) or Trichocereeae (i.e., *Brachycereus*), or being referred to a resurrected tribe Echinocereeae outside of the BCT clade.

Except for a few species of *Harrisia*, *Melocactus*, and *Pilosocereus*, all members of the BCT clade occur in tropical, subtropical, and temperate South America in a wide range of different habitats from coastal sand dunes and shrublands at sea level to more than 4000 m

in high-altitude grassland vegetation. This BCT clade contributes to a large extent to the two of the three cactus diversity centers in the Americas: (1) Altiplano and Andean foothills of Argentina, Bolivia, Chile, and Peru (primarily members of the tribe Trichocereeae) and (2) Caatinga and Campo Rupestre vegetations with tropical dry-deciduous forests and shrublands in northeastern Brazil (primarily members of the tribe Cereeae). The BCT clade includes a bewildering diversity of growth forms, ranging from tall and densely branched columnar cacti (e.g., *Browningia*, *Cereus*) to solitary large columnar and barrel-shaped forms (e.g., species of *Echinopsis*) to dwarf, globular plants consisting of single, unbranched and tuberculate stems only a few centimeters in diameter (e.g., species of *Rebutia*).

The two remaining tribes of the BCT clade are characterized primarily based on morphological characters of flowers and fruits, as well as biogeographical considerations (i.e., Buxbaum, 1969; Barthlott & Hunt, 1993; Anderson, 2001, 2005; Hunt & al, 2006). The tribe Cereeae sensu Hunt & al. (2006) consists of cacti that have the flowers arising primarily from lateral (i.e., older) areoles of their columnar stems or from a distinct fertile zone (i.e., cephalium) in either lateral or terminal position, with or without recurrent vegetative growth. The flowers are mostly regular, tubular, funnelform or salverform, and their floral segments are either whitish (i.e., flowers often nocturnal) or red, pink or rarely yellow (i.e., flowers diurnal). The pericarpel and floral tube is usually provided with scales but are otherwise naked. The fruits are fleshy, indehiscent or variously dehiscent, or only rarely almost dry. Genera assigned to the tribe Trichocereeae sensu Hunt & al. (2006) have the flowers arising laterally to subapically, and flowers are small to very large, and either white and nocturnal, or small to large and brightly colored, regular or more or less zygomorphic and diurnal. The pericarpel and floral tube is often covered by numerous scales and trichomes. The fruits are juicy or dry, and variously dehiscent. Earlier authors, such as Endler & Buxbaum (1974), characterized this tribe based on flowers whose pericarpel and floral tube have a concentration of usually numerous scales towards the base of the floral tube, and whose stamens are arranged in two distinct groups.

A close relationship between the members of the tribes Cereeae and Trichocereeae has been proposed for a long time (e.g., Buxbaum, 1969; Barthlott & Hunt, 1993), together with a rather close relationship to some similar-looking globular members of the tribe Notocacteae. In contrast, Barthlott (1979a, 1979b, 1988) favored a more broadly circumscribed taxon combining the taxa of Trichocereeae with those of Notocacteae and indicating close affinities of this combined taxon with the tribes Browningieae and Cereeae. Recent molecular phylogenetic studies confirmed some of these suspected relationships, and, hence, in a recent phylogenetic suprageneric classification of Cactaceae the circum-scription of the tribe Cereeae was expanded (i.e., tribe Cereeae s.l. sensu Nyffeler & Eggli [2010]) to coincide with the well supported BCT clade. Furthermore, a classification for the two major subclades at the subtribal rank was established (i.e., subtribes Cereinae and Trichocereinae; Nyffeler & Eggli, 2010). A third subtribe, Rebutiinae, was resurrected to include the genera of the basal grade in Cereeae s.l.

The assignment of several genera to either of the two tribes Cereeae and Trichocereeae, or to its closest relative, tribe Notocacteae, has been problematical and controversial, and different authors arrived at divergent solutions. This concerns the genera *Discocactus* (either seen to be closely related to *Gymnocalycium* or to *Melocactus*; Endler & Buxbaum, 1974), *Melocactus* (either included in Notocacteae or in Cereeae; Endler & Buxaum, 1974; Anderson, 2001, 2005), as well as *Facheiroa*, *Leocereus*, and *Uebelmannia*. Controversy also unfolded over a broader or narrower concept for genera like *Cleistocactus*, *Echinopsis*, and *Rebutia*. At the beginning 1980s an initiative to formulate a consensus for the previous disparate taxonomies of the current study group was set out (Hunt & Taylor, 1986, 1990), and in particular Hunt & al. (2006) favored wide circumscriptions of these genera, relegating previously separately kept genera, such as *Borzicactus*, *Lobivia*, or *Sulcorebutia*, to synonymy. In the case of *Echinopsis*, the history of favoring a conceptually wider circumscription dates back to Friedrich (1974a, 1974b) and Friedrich & Glätzle (1983).

We present here a comprehensive molecular phylogenetic analysis based on plastid DNA sequences using *trnK* intron (including *matK*), *rps16* intron, *rpl16* intron, and *trnS-trnG* intergenic spacer, in order to further resolve phylogenetic relationships among the major subclades and genera of the BCT clade (Nyffeler 2002; tribe Cereeae s.l. of Nyffeler & Eggli, 2010). Taxa of the tribe Trichocereeae sensu Hunt & al. (2006) are represented here with an expanded sampling in order to resolve taxonomic questions concerning their controversially discussed generic classification. Several phylogenetic studies have recently been published (e.g., Nyffeler, 2002; Crozier, 2005; Ritz & al., 2007; Meregalli & al., 2010; Demaio & al., 2011; Hernández-Hernández & al., 2011; Mosti & al., 2011, Schlumpberger & Renner, 2012) that address research questions concerning the circumscription of suprageneric taxa or selected genera (e.g., *Echinopsis*, *Gymnocalycium*, or *Rebutia*). However, these studies either applied a rather scattered taxon sampling not considering all relevant segregate genera, or used molecular markers rather variable to yield inferred sister-group relationship with adequate statistical support. In this study, we verify the phylogenetic status of the BCT clade as well as the circumscription and relationships of several of its subclades. In particular, we explore the phylogenetic and taxonomic status of the genera *Cleistocactus*, *Echinopsis*, *Espostoa*, and *Rebutia* as circumscribed by Hunt & al. (2006). We use statistical tests for the rejection of alternative topologies that are congruent to taxonomic concepts as favored by Hunt et al. (2006) in order to achieve well founded recommendations for adjustments in the generic classification without major disruptions. Furthermore, we are interested in exploring the effect of the interplay of more conserved (i.e., slow evolving) and more variable (i.e., fast evolving) genetic markers in order to resolve the backbone as well as, at the same time, the shallow relationships among the closely related species near the tips (e.g., Wiens & al., 2005). We aim for establishing a standard set of molecular markers from the plastid genome that might serve as standard for a broader and concerted effort to resolve relationships among, ultimately, all species of Cactaceae with the help of a combined data set in the form of a “supermatrix” (e.g., Fulton & Strobeck, 2006; Hinchliff & Roalson, 2013).

Material and Methods

Taxon sampling

For the present study we use the taxonomic concepts as outlined in Hunt et al. (2006): tribe Cereeae, tribe Trichocereeae, genera *Cleistocactus* s.l., *Echinopsis* s.l., *Espostoa* s.l., and *Rebutia* s.l. Hence, generic assignment and species classification of the exemplars follows Hunt & al. (2006). In contrast, our concept of tribe Cereeae s.l. includes all members of the BCT clade, and genera with narrower conceptions are referred to as *Cleistocactus* s.s., *Echinopsis* s.s., *Espostoa* s.s., and *Rebutia* s.s. In total, we sampled 74 ingroup exemplars (Table 1): 14 exemplars of the tribe Cereeae and 60 exemplars of the tribe Trichocereeae (Hunt & al., 2006). Furthermore, we added 16 exemplars as outgroups, including representatives of all former member genera of the tribe Browningieae (Table 1). For rooting the phylogenetic analyses we used *Calymmanthium substerile* and *Copiapoa bridgesii* as these two taxa they were found in previous phylogenetic studies (e.g., Nyffeler, 2002; Crozier, 2005; Bárcenas & al., 2011; Hérnandez-Hérnandez & al., 2011) to be among the most distantly related lineages within Cactoideae. Material for DNA extraction was mainly obtained from cultivated specimens grown at the Sukkulenten-Sammlung Zürich (Switzerland). A few samples were collected from herbarium specimens or from field collections (Table 2). Voucher specimens are deposited in the herbarium of the Sukkulenten-Sammung Zürich (ZSS).

DNA extraction, amplification, and sequencing

Total genomic DNA was isolated from silica dried tissue obtained from stems or flowers (receptacle and perianth segments) and extracted using DNeasy Plant Mini Kit (Qiagen, Basel, Switzerland). For taxa with high quantities of mucilage, the manufacturer's extraction protocol was modified as described in Nyffeler (2002). PCR reactions were performed in a total volume of 26 µL, containing 2.6 µL of 10x PCR buffer (Amersham Biosciences, Otelfingen, Switzerland), with 1.3 µL of 25 mmol/L $MgCl_2$, 3.0 µL of 1.25 mmol/L dNTPs, 0.8 µmol/L of corresponding primer, 1 unit of Taq DNA polymerase (Amersham Biosciences, Otelfingen, Switzerland) and 3 µL of genomic DNA. When the amplification was problematic, 1.0 µL of 5µg/µL bovine serum albumin (BSA, Sigma, Steinheim, Germany) and / or 1.0 µL of 5% DMSO was added. Amplification of the *trnK* intron, including the *matK* gene (primers trnK-3914F and trnK-2R; (Johnson & Soltis, 1994), and of the *rps16* intron (primers rpsF and rps2R; (Oxelman & al., 1997) were carried out with an initial denaturation step at 94°C (*trnK*/*matK*) and 95°C (*rps16* intron) for 4 min, followed by 39 cycles of melting at 94°C for 45 sec, annealing at 54°C (*trn*K/*mat*K) and 50°C (*rps16* intron) for 60 sec, an extension at 72°C for 90 sec and finished with a final elongation step at 72°C of 10 min. The *rpl16* intron was amplified using primers F71 (Jordan & al., 1996) and R1516 (Baum & al., 1998), and a protocol similar to previous one, but with the annealing temperature of 52°C for 45 sec and

an extension of 72°C for 2 min and 30 sec. The *trnS-trnG* intergenic spacer was amplified using primers trnS-F and trnG-R; (Hamilton, 1999) and applying the PCR protocol with a initial denaturation step of 95°C for 3 min, followed by 39 cycles of melting at 95°C for 3 min, annealing at 52° for 1 min, an extension at 72°C for 90 sec and an final elongation at 72°C of 10 min. Successfully amplified PCR products were purified using the GFX PCR DNA and a gel band purification kit (Amersham Biosciences, Otelfingen, Switzerland) according to the manufacturer's protocol and sequenced using Big Dye Terminator Cycle sequencing Reaction Kit (Perkin-Elmer, Applied Biosystems, Rotkreuz, Switzerland). Before loading on the automated sequencer ABI PRISM 3100 Genetic Analyzer (Perkin-Elmer), the reactions were purified on MicroSpin G-50 columns (Amersham Pharmacia Biotech Europe, Dübendorf, Switzerland) using multiscreen plates to remove excess Big Dye Terminator. Sequencing of both strands was accomplished by using the above mentioned external primers, and for trnK/matK, also by using six internal primers(trnK-23F, trnK-31R, trnK-41R, trnK-44F, trnK-52F, trnK-71R; (all from Nyffeler, 2002). Software Sequencher (version 4.2) (Gene Codes, Ann Arbor, Michigan, USA) was used to assemble complementary strands and individually check base positions for agreement.

Sequence alignment and phylogenetic analyses

Sequences were visually aligned using the software MacClade 4.0 (Maddison & Maddison, 2000). Despite numerous efforts to optimize PCR conditions, we were unable to generate sequences of all four gene regions for all taxa included in this study. Additionally, in some cases the sequences proofed to be unalignable with the rest of the matrix. This was especially the case for the *trnS-trnG* intergenic spacer marker for taxa from the outgroups. Therefore, those sequences were excluded from this study. As a result, our combined data set contains a number of "incomplete" exemplars, whose empty positions in the sequence matrix were specified as "missing data" and coded with quotation marks (following Wiens & al., 2005). All sequences were submitted to GenBank, and their accessions are shown in Table 2, and the aligned data set is available from Treebase [to be done].

Phylogenetic analyses were performed using maximum likelihood (ML) and maximum parsimony (MP) methods. Indels were not coded separately. For the ML analyses the four markers *trnK*/*matK*, *rpl16*, *rps16* and *trnS-trnG* were coded as eight separate partitions: 5' *trnK*, 1st and 2nd codon positions of *matK*, 3rd coding position of *matK*, 3'*trnK*, *rpl16*, *rps16*, and *trnS-trnG*. Preliminary analyses using PAUP* applied to the individual partitions did not yield any strongly supported incongruences, wherefore the eight different chloroplast partitions were combined. For both phylogenetic analyses methods we used PAUP*, version 4.0b10 (Swofford, 2002).

The maximum likelihood tree topology was calculated applying the GTR substitution model to all partitions, with all relevant model parameters estimated during the heuristic search. ML bootstrap support values were calculated with the software RAxML, version 7.3.2 (Stamatakis, 2006), applying 1'000 replicates of rapid bootstrap searches.

All most parsimonious trees were retained from a heuristic search employing 1'000 random addition sequence replicates and the tree bisection-reconnection (TBR) branch swapping algorithm, with STEEPEST and ALLSWAP options invoked. Relative support for each node was estimated using the bootstrap resampling procedure (Felsenstein, 1985) implemented in PAUP*, where the bootstrap values were calculated from 1'000 replicates, with 100 random addition replicates per bootstrap.

Comparing phylogenetic information content of molecular markers

We conducted separate MP analyses of the individual four aligned markers *trnK*/*matK*, *rpl16*, *rps16*, and *trnS*-*trnG* for the number of exemplars available. The first three analyses were rooted with *Calymmanthium substerile*, the *trnS*-*trnG* analysis was rooted with *Uebelmannia pectinifera*. For comparison, we tabulated the number of informative characters, consistency index (CI, excl. uninformative characters), retention index (RI, excl. uninformative characters), and the number of clades with MP bootstrap values higher than 50 percent for each of the four molecular markers.

Comparing phylogenetic information content of alternative hypotheses

We used parsimony- and likelihood-based paired-sites tests to investigate several null hypotheses that alternative tree topologies with monophyletic groups that represent specific taxonomic conceptions of earlier authors (i.e., Hunt et al., 2006) do not significantly differ from the best trees found in the maximum parsimony or maximum likelihood analyses. Only in the case of the rejection of alternative "constrained" topologies (i.e., traditionally used taxonomic concepts), we have a strong foundation to call for adjustments in the taxonomy (e.g., Nyffeler 2002). Constraints were defined for monophyletic groups and phylogenies were generated similar to outline above. Confidence limit was set to $P < 0.05$ with the enforced topological constraint. In a maximum parsimony framework we used the Wilcoxon signed-ranks test (i.e., Templeton test; Templeton, 1993) and the winning-sites test (Prager & Wilson, 1988), and in a maximum likelihood framework we used the Shimodaira-Hasegawa test (Shimodaira & Hasegawa, 1999; Goldman et al., 2000).

Results

Description of molecular markers and matrices

The four molecular markers were aligned in individual partitions and then concatenated into a combined cpDNA matrix (Table 3). Exemplars missing for certain markers were coded as missing information. Overall, the combined matrix consists of 5812 aligned nucleotide

positions, of which 578 proved to be informative for the total of 90 exemplars included in the matrix.

Phylogenetic analyses and inferred relationships

The ML and MP analyses yielded tree topologies that are illustrated in Figure 1 and Figure 2. Bootstrap support values from 1'000 replicates are mapped on the ML topology (log likelihood ln(L) = -23974.506) and on the strict consensus tree of 4061 MP topologies of length 2506. Overall, the inferred phylogenetic relationships are very similar in the two analyses. We find a strongly supported BCT clade (ML bootstrap support: 100%, MP bootstrap support: 100%) sister to exemplars of the tribe Notocacteae (i.e., genera *Eriosyce* and *Parodia*). Furthermore, the BCT clade (now referred to as tribe Cereeae s.l.) comprises three major subclades (i.e., identified with subtribal names in Figure 2; subtribe Cereinae, subtribe Rebutiinae, subtribe Trichocereinae) plus three distinct lineages corresponding to *Gymnocalycium*, part of *Rebutia* s.l. (clade R-1), and *Uebelmannia*. Both, *Rebutia* clade R-1 and *Uebelmannia* are part of a largely unresolved polytomy, with the remaining exemplars of Cereeae s.l. forming a moderately supported clade (78% ML, 79% MP). While the two subtribes Cereinae and Rebutiinae only receive moderate bootstrap support from both analyses (Cereinae: 80 ML, 65 MP; Rebutiinae: 88% ML, 73% MP), the support for Trichocereinae is very high (Trichocereinae: 100% ML, 100% MP). The largest subclade, representing in the present study by 45 exemplars that stand for a diversity of 15 genera based on the taxonomic concepts of Hunt & al. (2006) and some 280 species, fall into four well supported subclades (i.e., denominated CleWeb, DenLob, EchHar, and EspMat in Figure 2). *Arthrocereus* is found to be the sister-group to this group of subclades (56% ML, 53% MP). Finally, the four genera with broad circumscriptions (i.e., Hunt et al., 2006; *Cleistocactus* s.l., *Echinopsis* s.l., *Espostoa* s.l., and *Rebutia* s.l.) are not monophyletic based on findings from this study: *Cleistocactus* s.l. consists of five separate lineages (Figure 2; C-1 to C-5), *Echinopsis* s.l. consists of five lineages (Figure 2; E-1 to E-5), *Espostoa* s.l. consists of two lineages (Figure 2; S-1 and S-2), and *Rebutia* s.l. consists of two separate lineages (Figure 2; R-1 and R-2). Relationships within the subtribe Rebutiinae are found to differ comparing the ML and MP analyses: in the ML analysis, R-2 falls into two groups, with *Rebutia minuscula* (type species of *Rebutia* s.s.) sister to the other exemplars of the subtribe Rebutiinae. Furthermore, in the MP analysis, *Rebutia cintia* and *Rebutia steinbachii* are sister to each other, while in the ML analysis, *Rebutia cintia* is sister to *Rebutia fidana*. However, none of these deviating relationships receive distinct statistical support, indicating that they rather result from limitation in sequence data information to resolve them reliably.

Comparing phylogenetic information content of molecular markers

Molecular phylogenetic analyses addressing relationships over a deeper time period with a large group of species are often challenged by the limited capacity of molecular markers to provide comparative information to resolve deep as well as shallow relationships at the same time. The sampling strategy employed in this phylogenetic study follows "shortcut" approach as suggested by Wiens et al.(2005), consisting of simultaneously sampling "bottom-up" (i.e., deep relationships) and "top-down" (i.e., shallow relationships), which leads to incomplete character information for the exemplars in the matrix. In this study we included *trnK*/*matK* and *rps16* to contribute to resolve deep relationships among outgroup taxa, while *rpl16* and *trnS-trnG* were largely, or exclusively, sampled for ingroup exemplars. The strict consensus tree topologies derived from individual MP analyses of the four different markers as well as the combined cpDNA data set revealed that the fast evolving *trnS-trnG* intergenic spacers resolved 62 clades with 72 exemplars (20 MP topologies, CI excl. uninfor. char. = 0.40; RI = 0.69), while the other three markers were in this respect much less informative (data not shown). The combined data set, in comparison, resolved 60 clades with 74 exemplars (336 MP topologies, CI excl. uninfor. char. = 0.41; RI = 0.68). Inferred relationships between the two analyses are largely congruent. Hence, the trnS-trnG marker contributed significantly to resolving relationships in the tribe Cereeae s.l., but proofed to be too variable to be considered for deeper relationships. Furthermore, when bootstrap support values are considered for the major clades in Cereeae s.l. (Table 4) it becomes obvious that the expanded data set, including more conservative markers like *trnK*/*matK*, *rps16*, and *rpl16*, contribute information that improve the statistical support for inferred relationships. Bootstrap support values are distinctly (> 20%) higher for the combined data compared to the *trnS-trnG* data set for Cereinae, Rebutiinae, as well as CleWeb and, with a slightly lower difference, EspMat of Trichocereinae (Table 4). Furthermore, the number of clades identified with bootstrap values > 50% were distinctly larger for Cereinae as well as for EchHar and EspMat in Trichocereinae.

Comparing phylogenetic information content of alternative hypotheses

Parsimony- and likelihood-based paired-sites tests comparing constrained tree topologies representing alternative monophyletic groups were tested for 16 taxonomic concepts, using the Wilcoxon signed-ranks test (i.e., Templeton test; Templeton, 1983) and the winning-sites test in a maximum parsimony framework and the Shimodaira-Hasegawa test with full optimization (Shimodaira & Hasegawa, 1999; Goldman et al., 2000) in a maximum likelihood framework. Findings are listed in Table 5.

Discussion

Taxon sampling strategy and missing data

The sampling strategy in this study followed the sampling “shortcut” approach as suggested by Wiens & al. (2005) in order to resolve relationships among exemplar species over a wide range of the phylogeny. Sampling several slow-evolving molecular markers for a limited number of taxa contributes to resolving higher level relationships or the backbone (“scaffold”, Wiens 2006: 41) of the phylogeny. Furthermore, by adding additional exemplars sequenced for more rapidly evolving markers contributes to resolve shallow. For this study, outgroup exemplars were largely sampled for *trnK*/*matK* and *rps16*, while *rpl16* and *trnS*-*trnG* helped to resolve relationships among all currently recognized genera (e.g., Anderson, 2001, 2005; Hunt & al., 2006) and possible segregates. A consequence of this approach is a high amount of missing data in the combined data set. Still, comparing the inferred relationships from the *trnS*-*trnG* molecular data set (296 informative ingroup characters; Table 3) with the combined cpDNA data set (462 informative ingroup characters; Table 3) on the basis of the number of bootstrap support values for seven selected clades and the number of resolved clades included in them (Table 4) clearly illustrates the improved phylogenetic hypotheses derived from the combined data. For the subtribes Cereinae and Rebutiinae, the support value is distinctly higher, as is the case for two of the four subclades of Trichocereinae. Hence, the combined data managed to provide improved phylogenetic information for deep as well as shallow relationships.

Major relationships in the BCT clade and tribal and subtribal classification

The close relationships among the tribes Browningieae, Cereeae and Trichocereeae (e.g., Endler & Buxbaum 1974; Barthlott & Hunt, 1993) was first discovered by Nyffeler (2002, as "BCT clade"), and his findings were later corroborated by Crozier (2005), Bárcenas et al. (2011), Hernández-Hernández et al. (2011). The present study, again, provides for parsimony as well as likelihood analyses high support for the monophyly, and also confirms the tribe Notocacteae in recircumscribed form (Nyffeler & Eggli, 2010) as sister-group. Recently, Hunt & al. (2006) have incorporated taxa of former tribe Browningieae either in the tribe Echinocereeae (i.e., *Armatocereus*, *Jasminocereus*, *Neoraimondia*), or in the tribe Cereeae (i.e., *Browningia*, *Stetsonia*) or tribe Trichocereeae (i.e., *Brachycereus*). The tribes Cereeae and Trichocereeae sensu Hunt & al. (2006) are both found not monophyletic according to our analyses: Trichocereeae sensu Hunt & al. (2006) includes the genera *Leocereus*, *Facheiroa*, and *Discocactus*, which are found to form a clade in our analysis with

the members of a clade that includes the majority of Cereeae sensu Hunt & al. (2006), except for *Browningia* and *Uebelmannia*. The majority of Cereeae in Hunt's circumscription is not monophyletic (Table 5, Test 1). Likewise, Trichocereeae in Hunt's circumscription is firmly rejected on the basis of our constrained monophyly tests (Table 5, Test 2). Based on these strongly supported results and the support from several different investigations (e.g., Crozier, 2005; Bárcenas & al., 2011; Hérnandez-Hérnandez & al., 2011), we propose to characterize the tribe Cereeae in a circumscription expanded to include Trichocereeae sensu Hunt & al. (2006). This novel circumscription of Cereeae (referred to as Cereeae s.l. in our discussion) in a wide sense was already anticipated by Nyffeler & Eggli (2010), based on provisional data. Cereeae s.l. comprises 3 subclades comprising species of several different genera (independently also recovered by Schlumpberger & Renner [2012] but not identified as such) plus three lineages conforming to the genera *Gymnocalycium*, *Rebutia* sensu Hunt & al. (2006), and *Uebelmannia* (Figures 1 and 2). We propose to recognize these suprageneric subclades at subtribal level as Cereinae, Rebutiinae, and Trichocereinae.

Subtribe Cereinae includes the majority of genera of Cereeae sensu Hunt & al. (2006) and is unambiguously monophyletic with high support if *Browningia* is removed from the circumscription provided by Hunt & al. (2006) of this taxon. Cereinae in their majority have no spine- or hair-bearing areoles on their pericarpels and perianth tubes (notable exceptions are *Facheiroa* and the closely related *Leocereus* with their densely hairy flowers, which influenced their former placement in Trichocereeae), and the floral remains of many taxa rapidly turn black after anthesis. Cereinae are centered in the arid regions of NE Brazil and extend into the adjacent E Andean lowlands, and with a few species of *Pilosocereus* and *Melocactus* to N and NW South America, the Caribbean and S and SW Mexico.

Together with the observation that the basal sister of the tribe Cereeae, *Uebelmannia*, and the near-basal lineage *Arthrocereus* of subtribe Trichocereinae, both Brazilian endemics, a center of origin of the whole group in Brazil is likely. For both genera, the majority of species is confined to NE Brazil, but both have a couple of species in W South America (Peru, Ecuador, W Colombia), as well as a moderate number of species in N South America, throughout the Caribbean (with a modest radiation in Cuba in the case of *Melocactus*), and in SW and W Mexico. These parallel disjunctions to the W and to the N are notable in the light of the different pollination as well as fruit dispersal strategies (*Melocactus*: ornithophily/fruits likely consumed by lizards, small mammals and perhaps frugivorous birds; *Pilosocereus*: chiropterophily/ fruits likely consumed by bats and frugivorous birds, seed removal probably also by wasps). A similar case is the genus *Praecereus* (not included in our molecular study), where the single variable species *P. euchlorus* has a wide E Andean distribution but extends to W Andean slopes in Ecuador and Peru, and to N South America (Venezuela). All these cases could be examples of relatively recent long-distance dispersals.

Subtribe Rebutiinae is a clade with a novel circumscription, embracing *Browningia*, *Lasiocereus, Rebutia* s.s. and *Weingartia* s.l. (for *Rebutia* and *Weingartia* see expanded discussion below/above). The majority of taxa have flowers with naked axils of the pericarpel

and perianth scales, but *Lasiocereus* has densely hairy flowers, which apparently triggered its former placement in Trichocereeae. The exact topology of the four genera placed in subtribe Rebutiinae is still elusive. The detail analysis show *Rebutia* as basal sister to *Lasiocereus* and the other exemplars, and *Browningia* (two species analyzed) forms a grade in sister-group position to *Weingartia* s.l. Our statistical tests (Table 5, Test 7) could not reject monophyly of *Browningia*, but monophyly of a combined group including *Browningia* and *Lasiocereus* could also not be rejected (Table 5, Test 8).

Subtribe Trichocereinae is clearly monophyletic with very high support and conforms largely to Trichocereeae sensu Hunt & al. (2006), but excluding *Lasiocereus*, *Rebutia* sensu Hunt & al. (2006), *Facheiroa*, *Leocereus*, and *Gymnocalycium*. The majority of taxa have flowers with scant to abundant hairs in the axils of the scales on pericarpel and perianth tube. Other characters such as growth form, pollination syndromes, and fruit characteristics (i.e., juicy or dryish to dry, mode of dehiscence) are variable throughout the clade, although specific combinations of certain characters can be diagnostic for groups of species or genera. The very conservative approach to generic classification as favored by Hunt & al. (2006) led to the recognition of four large, polyphyletic genera. The circumscription of the genus *Echinopsis* is now for long controversially discussed, but also *Cleistocactus*, *Espostoa*, and *Rebutia* have been identified recently as heterogeneous and requiring resolution (e.g., Ritz & al. 2007).

Cleistocactus sensu Hunt & al. (2006)

The history of a wide circumscription of *Cleistocactus* dates back to the "consensus classification" reported by Hunt & Taylor (1986), where *Cleistocactus* basically united slender-growing columnar cacti with tubular, red and ornithophilous flowers. Hunt & Taylor (1986) commented that "unfortunately, *Borzicactus* sensu str. cannot be maintained as ... distinct from *Cleistocactus*; the floral differences ... are inadequate". In addition, they remarked that *Oreocereus*, including *Matucana* and *Oroya*, with a similar flower syndrome in most species, could only "perhaps" be maintained at generic level. Our tests of the monophyly of *Cleistocactus* sensu Hunt & al. (2006) clearly reject such a circumscription (Table 5, Test 3) and the concept is thus not tenable and must be refused as being polyphyletic. The species of *Cleistocactus* sensu Hunt & al. (2006) are dispersed over four to five different lineages in the CleWeb and EspMat subclades of Trichocereinae (Figure 2). Clade C-1 is a small group of prostrate to hanging, slender, cereoid plants with zygomorphic ornithophilous flowers from the foothills of the Bolivian Andes, in sister-group position to the presumably chiropterophilous *Samaipaticereus corroanus*, from the same general area in Bolivia. The earliest name of this segregate is *Bolivicereus*, which we propose to accept at generic rank with the three species *Bolivicereus samaipatanus, B. aureispinus* and *B. colademononis*.

Clade C-2 includes the type of the genus and thus conforms to *Cleistocactus* s.s. - cereoid, usually upright-growing plants with tubular, actinomorphic or zygomorphic,

ornithophilous flowers. The close association of *Cleistocactus* s.s. with *Vatricania* (see discussion below for *Espostoa* s.l.), *Weberbauerocereus* and *Yungasocereus* is notable. Ritter (1980) and Mottram (2006) already assumed a close relationship of *Yungasocereus inquisiviensis* with *Cleistocactus* s.s. The overall similarity of the allegedly chiropterophilous flowers of *Yungasocereus* with the ornithophilous flowers of *Cleistocactus* s.s. (esp. *C. laniceps*, which is also vegetatively very similar to *Yungasocereus*, see illustrations in Mottram, 2006), is notable. His conclusion that the clade that includes *Samaipaticereus* and *Yungasocereus* should be recognized as a vastly expanded *Espostoa*, that would also include the Brazilian endemic *Facheiroa*, is, however, not supported to any degree.
Clade C-3 conforms to *Borzicactus* in the original sense (but not sensu Kimnach (1960), see below) with prostrate-erect to erect, columnar growth and tubular, zygomorphic, ornithophilous flowers.

Clades C-4 and C-5 represent two species that are part of a polytomy that also includes exemplars of the genera *Haageocereus, Oreocereus,* and *Matucana*. Our data is inconclusive whether the two taxa included in our sampling form a single clade or two separate clades, and about the possible relationships with *Haageocereus*. In habit and floral characteristics, clade C-4 resembles C-3, but it should be noted that some *Haageocereus* species also have red to reddish, somewhat zygomorphic, ornithophilous flowers (e.g., *Haageocereus pseudomelanostele* ssp. *carminiflorus*; Calderón & al. 2007). Monophyly of a taxon including C-4 and C-5 was not rejected by our three tests (Table 5, Test 11), and it could be recognized at generic level under the name *Loxanthocereus*.
When discussing *Borzicactus*, the earlier broad concept of Kimnach (1960) must also be briefly considered: *Borzicactus* s.l. sensu Kimnach (1960) was diagnostically characterized by zygomorphic ornithophilous flowers, but all statistical tests we employed clearly rejected this broad circumscription (Table 5, Test 9). Even when the strictly Bolivian *Bolivicereus* is excluded, *Borzicactus* s.l. is rejected (Table 5, Test 10). *Borzicactus* s.l. sensu Kimnach (1960) is thus yet another example of placing overemphasis on floral characters (i.e., zygomorphic ornithophilous flowers), resulting in the inclusion of several disparate elements, including *Matucana*, *Oreocereus* (incl. *Arequipa* and *Morawetzia*), *Bolivicereus* (to be recognized at generic rank according to our data), *Loxanthocereus* (probably to be recognized at generic rank according to our data), but not *Oroya*, because this genus has actinomorphic, shortly tubular-urceolate flowers.

Echinopsis sensu Hunt & al. (2006)

The wide and equally or even more controversial circumscription of the genus *Echinopsis* dates back to Friedrich (1974a), Friedrich (1974b) and Friedrich & Glätzle (1983), although these authors still recognized *Lobivia* as a separate entity. A very wide circumscription of *Echinopsis* was supported by the consensus classification as reported by Hunt & Taylor (1986), who included also *Lobivia*, remarking that even though especially the loss of the traditional genus *Lobivia* is "regrettable" but "seems unavoidable from both taxonomic and

purely practical considerations". Prior to the consensus classification attempts, the components that were to form *Echinopsis* s.l. were separated from each other primarily on the basis of growth form (columnar, barrel-shaped, dwarf globose) and floral characters (i.e, time of anthesis and coloration, which primarily represent the pollination syndromes of sphingophily and melittophily). Our analysis unequivocally shows that this wide circumscription of *Echinopsis* is rejected by all our statistical tests employed (Table 5, Test 4) and such a wide circumscription results in a completely artificial polyphyletic taxon.

Our analyses show that *Echinopsis* sensu Hunt & al. (2006) falls in three distinct groups, each of which is part of a subclade that involves other, widely accepted separate genera. The much more inclusive sampling of the genus *Echinopsis* by Schlumpberger & Renner (2012) largely corroborates our inferred relationships. Hence, these two studies provide insights on the diversity and relationships among the different subgroups. While, originally, these authors suggested to wait with formal taxonomic changes, a subsequent paper by Schlumpberger (2012), and supported by Hunt (2012), provided the necessary species combinations for the suggested circumscriptions of monophyletic segregate genera. We largely follow this line of reasoning, but remain more conservative in only recognizing segregates that are either monophyletic based on the present investigations (Figures 1 and 2), or whose potential monophyly cannot be rejected based on the present data. *Echinopsis* s.s., as suggested by us, consists of the two clades E-1 and E-2 (Figure 2; Table 5, Test 12), but was split by Schlumpberger & Renner (2012), based on largely similar inferred relationships into two genera, *Echinopsis* and *Leucostele* (Schlumpberger, 2012). Similarly, we recognize *Lobivia* as distinct (clade E-5 in Figure 2), but including *Echinopsis silvestrii* (= *Chamaecereus silvestris*) and *Echinopsis tarijensis* (= *Soerensia tarijensis*). Furthermore, we recognize *Trichocereus* in the narrow circumscription of early authors (e.g., Backeberg, 1966). The recent paper by Albesiano & Terrazas (2012), investigating the limits of *Trichocereus*, does not provide additional data due to inadequate outgroup taxon sampling. Friedrich & Glätzle (1983) were early proponents of a widely circumscribed *Echinopsis*, though still recognizing *Lobivia* as separate. Based on their study of seed characters concerning outline shape, form of hilum, micromorphology of the testa, they proposed possible lines of subgeneric divisions of *Echinopsis* s.l. However, their seed characters do not parallel clades identified in our study as well as the one of Schlumpberger & Renner (2012).

Espostoa sensu Hunt & al. (2006)

The wider circumscription of *Espostoa*, including the monotypic genera *Thrixanthocereus* and *Vatricania*, also dates back to the consensus approach reported by Hunt & Taylor (1986). The widening of the circumscription was prompted by the common character of nocturnal chiropterophilous flowers appearing from a lateral woolly to bristly cephalium. Our tests for monophyly of this expanded circumscription are all rejected such a circumscription (Table 5, Test 5), and *Espostoa guentheri* must be excluded and should be accepted as

monotypic genus *Vatricania*. *Vatricania* is shown as sister of *Cleistocactus* s.s. in the CleWeb clade in our consensus phylogeny (Figure 2), while the clade of the remaining *Espostoa* taxa is part of the EspMat clade. This finding on sister-group relationships of *Espostoa* s.s. and *Vatricania* largely correspond to that formulated by Ritter (1980) on morphological grounds. The remaining taxa of *Espostoa* s.s. form a monophyletic clade that also includes the monotypic genus *Rauhocereus* (chiropterophilous flower, no cephalium). While the tree topology (Figure 1) shows *Rauhocereus* embedded in *Espostoa* s.s., our statistical tests (Table 5, Test 13) did not reject *Espostoa* s.s. excluding *Rauhocereus*. Our sampling is thus insufficient for a final decision, and additional taxa of *Espostoa* must be analyzed to obtain clarity about the proper placement of the enigmatic Peruvian genus *Rauhocereus*. The close relationship of *Rauhocereus* with *Espostoa* is further supported by the existence of a natural hybrid with *Espostoa* (Ritter 1981: 1344), although the author, also, assumed a close relationship between *Rauhocereus* and *Weberbauerocereus*.

Rebutia sensu Hunt & al. (2006)

The wide and controversial circumscription of *Rebutia* to include *Sulcorebutia* and *Weingartia,* as well as *Aylostera, Digitorebutia* and *Mediolobivia*, dates back to the consensus classification summarized by Hunt & Taylor (1986), who recognized that the taxon is "a mixture of convergent forms and not monophyletic". The unifying characters for this wide circumscription of the genus where the growth form (i.e., tuberculate, small globose, solitary to offsetting plants) and the brightly colored diurnal actinomorphic flowers, with or without hairs and/or spines in the axils of the pericarpel and perianth scales. Opinions amongst the participants of the consensus process were mixed, however, and even an inclusion of *Rebutia* s.l. into a much enlarged *Echinopsis* was considered. In our analysis, monophyly of *Rebutia* sensu Hunt & al. (2006) could be rejected with absolute certainty (Table 5, Test 6). We find that components of *Rebutia* s.l. are placed in two separate major lineages of Cereeae s.l., conforming to the results of Ritz & al. (2007, with much deeper sampling of *Rebutia* s.l.). Their clade "*Rebutia* I" (our clade R-1 in Figure 2) includes taxa of the formerly segregate genera *Aylostera, Mediolobivia* and *Digitorebutia*, and their second clade includes taxa of *Rebutia* s.s. ("*Rebutia* II") plus all species of *Sulcorebutia, Cintia* and *Weingartia* (our clade R-2 in Fig. 2). Due to the sampling employed, Ritz & al. (2007) could not unambiguously resolve the relationships within Cereeae s.l. However, there is complete agreement between our results and those of Ritz & al. (2007) that clade R-2, which include taxa formerly assigned to *Sulcorebutia, Cintia* and *Weingartia*, form a single clade, for which the name *Weingartia* has to be used for nomenclatural priority reasons. Finally, the placement of *Rebutia* s.ss. remains ambiguous on the base of the available data: Ritz et al. (2007) place these species (their "*Rebutia* II" clade) in sister-group position to *Browningia*, while our analyses are inconclusive concerning relationships of the type species of *Rebutia*.

Relationships within subtribe Trichocereinae and recognition of four suprageneric subclades (plus Arthrocereus)

Within our subtribe Trichocereinae, the Brazilian endemic cereoid genus *Arthrocereus* occupies a sister-group position to all other taxa. The support for the formation of all other taxa of Trichocereinae to form a clade sister to *Arthrocereus* is moderate to high (ML 59%, MP 96%; Figs. 1 and 2). Four subclades are further recognized here within Trichocereinae:

The first subclade, EchHar (from *Echinopsis* and *Harrisia*), which includes the type of the genus *Echinopsis* and thus *Echinopsis* s.s., as well as columnar taxa presently placed in *Echinopsis* s.l. sensu Hunt & al. (2006) and recognized as separate genus *Leucostele* by Schlumpberger (2012), plus the genus *Harrisia*. The biogeography of the genus *Harrisia* is puzzling. The genus embraces 5 species from the E-Andean lowlands (including 1 from NE Brazil) and 14 species from the Caribbean region. The genus is unquestionably monophyletic (corroborated by Schlumpberger & Renner 2012 and Franck & al. 2013), and most enigmatically, the ancestral lineages within the genus are Caribbean, and the derived taxa S American. The present-day geographical range of *Harrisia* suggests two long-distance distribution events as inferred by Franck & al. (2013). Furthermore, the Leucostele group within *Echinopsis* is also biogeographically interesting, as it shows a Trans-Andean occurrence.

The second subclade, CleWeb (from *Cleistocactus* and *Weberbauerocereus*), is formed by taxa the of *Samaipaticereus, Yungasocereus,* part of *Cleistocactus* s.l. sensu Hunt & al. (2006), *Weberbauerocereus*, and part of *Espostoa* s.l. sensu Hunt & al. (2006). All taxa are cereoid in growth form. The origin of this clade is likely E-Andean (*Yungasocereus*, *Vatricania* and *Cleistocactus* s.s. all occur on the E Andean slopes), and only *Weberbauerocereus* has, with the exception of a recently described species from Bolivia, a W-Andean occurrence.

The third subclade, DenLob (from *Denmoza* and *Lobivia*), is again very strongly supported, and apart from taxa of *Denmoza* and *Acanthocalycium* also includes numerous species of *Echinopsis* sensu Hunt & al. (2006), which are referred to the resurrected genus *Lobivia*. The majority of the taxa of this subclade have a globose growth form, but Trichocereus are generally columnar, and some *Lobivia* are barrel-shaped. The origin of the DenLob subclade appears to be W Andean, as the strictly W-Andean *Trichocereus* s.s. as identified as the most ancestral lineage, followed by strictly E-Andean clades (pointing to an evolutionary ancient disjunction). Within these E-Andean clades, one species each of *Lobivia* and *Echinopsis* have extended their modern distributions to the upper altitudes of the W Andean slopes (pointing to an evolutionary young range extension).

The forth subclade, EspMat (from *Espostoa* and *Matucana*), is likewise strongly supported and includes part of *Espostoa* s.l. sensu Hunt & al. (2006), *Rauhocereus*, part of *Cleistocactus* sensu Hunt & al. (2006), *Haageocereus*, *Mila*, *Pygmaeocereus*, *Oreocereus*, *Matucana* and *Oroya*, with various growth forms varying from small and globular to cereoid and tree-like. The consensus classification reported on by Hunt & Taylor (1986) was undecided about the circumscription of *Haageocereus*, and it was speculated that

Pygmaeocereus could be closely related. This is confirmed by our data, and *Pygmaeocereus* as well as *Mila* (unambiguously accepted at generic rank by Hunt & Taylor 1986, 1990) are placed in a polytomy with *Haageocereus* in the consensus phylogeny (Figure 2), while the detail analyses show *Mila* and *Pygmaeocereus* in a derived position within a clade that includes *Haageocereus* and *Loxanthocereus* (see discussion above under *Cleistocactus*). However, our statistical tests are all unable to reject an expanded genus *Haageocereus* (i.e. including *Mila* and *Pygmaeocereus*, as suggested on the base of preliminary data by Nyffeler & Eggli 2010; Table 5, Test 14). Again, more data is needed to elaborate the topology of the whole clade. The consensus classification reported on by Hunt & Taylor (1986) included *Matucana* and *Oroya* in *Oreocereus*, but uncertainty was formulated whether *Oreocereus* s.l. could be accepted as separate from the expanded genus *Cleistocactus* s.l. (= sensu Hunt et al. 2006). Our consensus phylogeny (Figure 2) shows *Oroya* (actinomorphic flowers) imbedded in *Matucana* (zygomorphic ornithophilous flowers), but support for this configuration is low: the three statistical tests employed (Table 5, Test 15) did not reject monophyly for *Matucana* s.s., while exclusion of *Oroya* from *Matucana* s.s. was only rejected by one out of three tests (Table 5, Test 16). Available data is thus inconclusive as to the placement of *Oroya*. The origins of the EspMat subclade also appears to be W Andean, as the majority of the subclade has a primarily W Andean occurrence, though with many taxa in the river Marañón drainage (ultimately a tributary of the Amazonas river). Only *Oreocereus* has taxa on the E slopes of the Andes, likely pointing to an evolutionary young range extension.

Orphans

Aylostera/*Mediolobivia*: discussed above.
Gymnocalycium: The genus *Gymnocalycium*, comprising some 80 to 90 species, is identified in these analyses as sister-group to the tribe Trichocereinae. The phylogenetic position of *Gymnocalycium* was controversially discussed for a long time: tt was included in the tribe Notocacteae by Endler & Buxbaum (1974), but in Trichocereeae by Barthlott & Hunt (1993), Anderson (2001, 2005), Hunt & al. (2006), and Ritz & al. (2007). Its placement as part of the basal grade of genera of Cereeae s.l. is, however, clearly supported and independently corroborated by Schlumpberger & Renner (2012), even though the exact topology remains ambiguous.
Uebelmannia: At the base of Cereeae s.l., the genera *Uebelmannia* and a clade of species that include the type species of *Aylostera* and *Mediolobivia* form a basal grade or a polytomy with a clade including the other species of Cereeae s.l.

The genera of the basal grade of Cereeae (i.e., *Uebelmannia*, *Aylostera*, and *Gymnocalycium*), as well as all genera of Cereinae, and the genera of the basal grade of Trichocereinae (*Reicheocactus* [see Schlumpberger & Renner, 2012; not analyzed by us], *Arthrocereus*) have an exclusively E Andean distribution (except *Pilosocereus* and *Melocactus*, see above).

Growth form evolution

Growth forms vary enormously in Cereeae s.l. as here circumscribed, from dwarf globose forms to large arborescent forms, and from solitary unbranched to shrubby or crown-forming taxa. Shared growth forms were commonly used in the past either as diagnostic characters to identify relationships or to circumscribe taxa (such as the segregation of *Lobivia*, etc. from *Echinopsis* and *Trichocereus*, see above). Both approaches fail because of the numerous parallel shifts from one growth form to another, as revealed by the topology of our phylogeny (Figures 1, 2): formal ancestral state reconstruction has not been performed, but all evidence points to small globose growth forms as being the ancestral state for Cereeae - both its immediate sister-group (tribe Notocacteae) as well as all three genera of the basal grade belong to this growth form type. For subtribe Rebutiinae, globose growth again appears to represent the ancestral condition, while for subtribe Cereinae, columnar growth is likely the ancestral condition. This implies a propensity for transitions from globose to columnar. In both subtribes, a second transition back from columnar to globose appears likely, but our topologies are not sufficiently resolved for definite statements.

For Trichocereinae, the situation appears complex and puzzling in the light of the well-supported discovery by Schlumpberger & Renner (2012) that the unbranched dwarf globose taxa of *Reicheocactus* (not analyzed by us, and part of *Echinopsis* s.l. sensu Hunt & al. (2006)) represent the basal sister clade to the remaining subtribe. The most basal emerging lineage of the remaining subtribe is likely cereoid-shrubby *Arthrocereus*, recovered in a trichotomy with EchHar and AcaDen+EspHag subclades in our consensus phylogeny (in maximum parsimony analyses, Figure 1), or as basal sister to the subclades just named in Bayesian analyses. Within the four named subclades, it appears that cereoid growth form might represent the ancestral condition for DenLob and EspMat, with secondary transitions to globose forms in either case. For the EchHar subclade, no statement as to growth form evolution is possible, and the members of the CleWeb clade are all cereoid. It thus becomes amply clear that there are numerous parallel transitions between growth forms throughout the subtribe, with no universal trend in either direction apart from the observations that ancestral lineages are more commonly small and globose in growth form, and that cereoid subclades (at least DenLob and EspMat) likely have some more advanced lineages with a transition to globose growth again. Overall, growth form is a fairly labile character and thus in a broad usage unsuitable for recognizing evolutionary units. This observation also extends to the evolution of cephalia, which evolved repeatedly in parallel in both Cereinae (at least three separate origins involving several genera, lateral or terminal) and Trichocereinae (three separate origins for *Cephalocleistocactus* [not studied by us but shown as sister to *Yungasocereus* by Schlumpberger & Renner 2012], *Vatricania* and *Espostoa*).

Synopsis of the tribe Cereeae and new combinations

Family **Cactaceae** Juss. (1789)
Subfamily **Cactoideae** Eaton (1836)
Tribe **Cereeae** Salm-Dyck (1840)

Genera unassigned to subtribe: ***Aylostera*** (R-1; potentially *Mediolobivia* as separate genus), ***Gymnocalycium***, ***Uebelmannia***

Subtribe **Cereinae** Britton & Rose (1920)
Accepted genera: ***Cereus, Stetsonia, Cipocereus, Brasilicereus, Leocereus, Facheiroa, Micranthocereus, Espostoopsis, Coleocephalocereus Discocactus, Melocactus, Arrojadoa, Pilosocereus***

Subtribe **Rebutiinae** Donald (1955)
Accepted genera: ***Browningia, Lasiocereus, Rebutia, Weingartia*** (potentially *Cintia* and/or *Sulcorebutia* as separate genera as well)

Subtribe **Trichocereinae** Buxbaum (1958)
Accepted genera: ***Arthrocereus***, and an additional 24 genera grouped into four clades

Clade I (CleWeb): ***Bolivicereus*** (C-1), ***Cleistocactus*** (C-2), ***Cephalocleistocactus*** (found to be "distinct" from Cleistocactus by Schlumpberger & Renner, 2012), ***Samaipaticereus***, ***Vatricania***, ***Weberbauerocereus***, ***Yungasocereus***

Clade II (EspMat): ***Borzicactus***, ***Espostoa***, ***Haageocereus***, ***Loxanthocereus***, ***Matucana****, ***Mila***, ***Oreocereus***, ***Oroya***, ***Rauhocereus*** (potentially to be included in *Espostoa*, further investigations to reach highly supported result needed), ***Pygmaeocereus***.
* several variants for combining *Matucana*, *Oreocereus*, *Loxanthocereus* (C-4 and C-5), *Haageocereus*, *Pygmaeocereus* and *Mila* are possible, further investigations to reach well supported result needed.

Clade III (DenLob): ***Acanthocalycium***, ***Denmoza***, ***Lobivia***** (potentially with distinct *Chamaecereus*, *Soehrensia*), ***Setiechinopsis***, ***Trichocereus***
** *Chamaecereus* and *Lobivia* both have been described by Britton & Rose on Oct. 12, 1922, *Soehrensia* by Backeberg in 1938.

Clade IV: (EchHar): ***Echinopsis*** (potentially *Leucostele* as separate genus, further investigations to reach highly supported result needed), ***Harrisia***

Taxonomic changes:

Bolivicereus aureispinus (F. Ritter) Nyffeler & Eggli, comb. nov. Basionym: *Winteria aureispina* F. Ritter, Kakt. and. Sukk. 13: 4, 1962. Nomenclatural note: The name *Winteria aureispina* is incorrect as *Winteria* F. Ritter 1962 is a later illegitimate homonym of *Wintera* Murry 1784, but it is available as basionym.
Lobivia tarijensis (Vaupel) Nyffeler & Eggli, comb. nov. Basionym: *Cereus tarijensis* Vaupel, Monatsschrift für Kakteenkunde 26(8): 123–124.

Conclusions

1. The tribes Cereeae and Trichocereeae as recognized by Hunt & al. (2006) should be combined to form an expanded tribe Cereeae s.l., conforming to the BCT clade as originally defined by Nyffeler (2002) and later confirmed by several studies. Cereeae s.l. is sister to the tribe Notocacteae, which in recircumscribed form comprises the genera *Eriosyce*, *Neowerdermannia*, *Parodia*, *Rimacactus*, and *Yavia* (Nyffeler & al. unpubl.).

2. Tribe Cereeae as defined here, consists of three distinct lineages recognized as genera (i.e., *Uebelmannia*, *Aylostera*, *Gymnocalycium*) and three subclades which we recognize as subtribes Cereinae, Rebutiinae, and Trichocereinae.

3. Subtribe Cereinae largely corresponds to the tribe Cereeae sensu Hunt & al. (2006), including *Discocactus*, *Facheiroa*, and *Leocereus*, but excluding *Browningia* and *Uebelmannia*.

4. Subtribe Rebutiinae is a newly circumscribed clade that comprises the genera *Lasiocereus, Browningia, Rebutia*, and *Weingartia* s.l. (potentially with distinct *Cintia* and/or *Sulcorebutia*). It differs from the circumscription of Nyffeler & Eggli (2010) by being restricted to monophyly, excluding *Aylostera*, *Gymnocalycium*, *Stetsonia* (now included in Cereinae) and *Uebelmannia*.

5. Subtribe Trichocerinae is to be recircumscribed based on the tribe Trichocereeae sensu Hunt & al. (2006) by excluding *Discocactus*, *Facheiroa*, and *Leocereus* (now Cereinae), *Gymnocalycium*, *Lasiocereus*, and *Rebutia* s.l. (now either Rebutiinae or *incertae sedis*).

6. *Uebelmannia* (three morphologically variable species, endemic to NE Brazil) is identified as representing one of the cladistically basal lineages of Cereeae s.l.

7. The widely circumscribed genera *Cleistocactus, Echinopsis, Espostoa* and *Rebutia* in the sense of Hunt & al. (2006) are all polyphyletic and include two or more distinct lineages.

8. *Cleistocactus* sensu Hunt & al. (2006) consists of three distinct lineages recognized as *Bolivicereus, Cleistocactus* s.s. and *Borzicactus*. The phylogenetic position of the remaining lineages (*Loxanthocereus*) remains ambiguous at present.

9. *Echinopsis* sensu Hunt & al. (2006) is scattered among two distinct subclades of Trichocereinae (i.e., EchHar and DenLob) and we propose to recognize at least *Echinopsis* s.s., *Trichocereus*, *Lobivia*, as well as *Acanthocalycium* and *Setiechinopsis*.

10. *Espostoa* s.l. sensu Hunt & al. (2006) consists of two different clades, and the monotypic genus *Vatricania* ought to be recognized at generic level.

11. *Rebutia* s.l. sensu Hunt & al. (2006) consists of 2 distinct lineages: *Aylostera/Mediolobivia* (incl. *Digitorebutia*) is part of the basal lineages of Cereeae s.l., while *Rebutia* s.s. and *Weingartia* (incl. *Sulcorebutia* and *Cintia*) are placed in subtribe Rebutiinae. Whether *Rebutia* s.s. and *Weingartia* s.l. represent two separate lineages, or should be combined into one genus, remains ambiguous at present.

12. Tribe Cereeae s.l. is basically a South American taxon, with only few species of three genera (*Harrisia, Melocactus, Pilosocereus*) extending to the Caribbean and S Mexico. Several East/West Andean transitions are recognized, and both, ancient long-distance dispersals as well as younger range expansion events have likely occurred.

Suggestions for future research

The topology of the EspMat clade is still unsatisfactorily resolved, and a deeper sampling of the genera here placed should be attempted in order to get a clear picture of growth form transitions. Also the likely paraphyly of *Espostoa* versus *Rauhocereus* remains to be corroborated. More detailed sampling of the genera *Browningia* and *Weberbauerocereus* would allow the formulation of additional biogeographic hypotheses. Finally, a more detailed sampling involving additional species of *Rebutia* s.s. should be envisaged in order to clarify the position of *Rebutia* s.s. versus *Weingartia*. The same applies to *Aylostera* and *Mediolobivia*.

References

Albesiano, S. & Terrazas, T. 2012. Cladistic analysis of *Trichocereus* (Cactaceae: Cactoideae: Trichocereeae) based on morphological data and chloroplast DNA sequences. Haseltonia 17: 3-23.

Anderson, E.F. 2001. The Cactus Family. Portland, Oregon, USA: Timber Press.

Anderson, E.F. (transl. Eggli, U.). 2005. Das grosse Kakteen Lexikon. Stuttgart: Ulmer Verlag.

Backeberg, C. 1958-1962. Die Cactaceae. 6 Vols. Jena: G. Fischer Verlag.

Backeberg, C. 1966. Das Kakteenlexikon. Jena: G. Fischer Verlag.

Bárcenas, R., Yesson, C. & Hawkins, J.A. 2011. Molecular systematics of the Cactaceae. Cladistics 27: 1-20.

Barthlott, W. 1979. Cacti. Cheltenham, GB: Stanley Thornes Ltd.

Barthlott, W. & Voit, G. 1979. Mikromorphologie der Samenschalen und Taxonomie der Cactaceae: Ein raster-elektronenmikroskopischer Überblick. Pl. Syst. Evol. 132: 205-229.

Barthlott, W. 1988. Über die systematische Gliederung der Cactaceae. Beitr. Biol. Pfl. 63: 17-40.

Barthlott, W. & Hunt, D.R. 1993. Cactaceae. In: Kubitzki, K. et al. (eds.). The families and genera of vascular plants, Vol. 2, pp. 161-197. Berlin: Springer Verlag.

Baum, D.A., Small, R.L. & Wendel, J.F. 1998. Biogeography and floral evolution of baobabs (Adansonia, Bombacaceae) as inferred from multiple data sets. Syst. Biol. 47: 181-207.

Buxbaum, F. 1958. The phylogenetic division of the subfamily Cereoideae, Cactaceae. Madroño 14: 177-206.

Buxbaum, F. 1968. Die Entwicklungslinien der Tribus Cereeae Britt. et Rose emend. F. Buxbaum (Cactaceae-Cactoideae). II. Versuch einer phylogenetischen Gliederung innerhalb der Tribus. - Beitr. Biol. Pfl. 44: 389-433.

Buxbaum, F. 1969. Die Entwicklungslinien der Kakteen in Südamerika. In: Fittkau, E.J. et al. (eds.), Biogeography and ecology in South America, Vol 2, pp. 583-623. Den Haag, NL: Dr. W. Junk.

Buxbaum, F. 1975. Provisorium einer Neugliederung der Tribus Trichocereae(CV). In: Krainz, H. (ed.), Die Kakteen, CV. 1 Bl. [1. 1. 1975].

Calderón, N., Zappi, D., Taylor, N. & Ceroni, A. 2007. Taxonomy and conservation of Haageocereus Backeb. (Cactaceae) in Peru. Bradleya 25: 45-124.

Crozier, B.S. 2005. Systematics of Cactaceae Juss.: Phylogeny, cpDNA evolution, and classification, with emphasis on the genus Mammillaria. Austin, TX, USA: Dissertation, University of Texas at Austin.

Demaio, P.H., Barfuss, M.H., Kiesling, R., Till, W. & Chiapella, J.O. 2011. Molecular phylogeny of Gymnocalycium (Cactaceae): assessment of alternative infrageneric

systems, a new subgenus, and trends in the evolution of the genus. Amer. J. Bot. 98: 1841-1854.

Edwards, E.J., Nyffeler, R. & Donoghue, M.J. 2005. Basal cactus phylogeny: implications of *Pereskia* (Cactaceae) paraphyly for the transition to the cactus life form. Amer. J. Bot. 92: 1177-1188.

Endler, J. & Buxbaum, F. 1974. Die Pflanzenfamilie der Kakteen. 3rd ed. Minden, D: Lehrmeister-Bücherei Nr. 872.

Felsenstein, J. 1985. Confidence limits on phylogenies - an approach using the bootstrap. Evolution 39: 783-791.

Franck, A.R., Cochrane, B.J. & Garey, J.R. 2013. Phylogeny, biogeography, and infrageneric classification of *Harrisia* (Cactaceae). Syst. Bot. 38: 210-223.

Friedrich, H. 1974a. Zur Taxonomie und Phylogenie der Echinopsidinae (Trichocereinae). IOS Bull. 3: 79-99.

Friedrich, H, 1974b. *Lobivia* oder *Echinopsis*? Kakt. and. Sukk. 25: 55-57, 80-83.

Friedrich, H. & Glätzle, W. 1983. Seed-morphology as an aid to classifying the genus *Echinopsis* Zucc. Bradleya 1:

Fulton, T. L. & C. Strobeck. C. 2006. Molecular phylogeny of the Arctoidea (Carnivora): effect of missing data on supertree and supermatrix analyses of multiple gene data sets. Mol. Phyl. Evol. 41: 165-181.

Gibson, A.C. & Nobel, P.S. 1986. The Cactus Primer. Cambridge, MA, USA: Harvard University Press.

Goldman, N., Anderson, J.P. & Rodrigo, A.G. 2000. Likelihood-based tests of topologies in phylogenetics. Syst. Biol. 49: 652-670.

Hamilton, M.B. 1999. Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. Mol. Ecol. 8: 521-523.

Hérnandez-Hérnandez, T., Hérnandez, H.M., De-Nova, J.A., Puente, R., Eguiarte, L.E. & Magallon, S. 2011. Phylogenetic relationships and evolution of growth form in Cactaceae (Caryophyllales, Eudicotyledoneae). Amer. J. Bot. 98: 44-61.

Hinchliff, C.E. & Roalson, E.H. 2013. Using supermatrices for phylogenetic inquiry: an example using the sedges. Syst. Biol. 62: 205-219.

Hunt, D.R. (ed.). 2006. The New Cactus Lexicon.

Hunt, D.R. 2012. Tricky Trichos. Cact. Syst. Init. 28: 3-4.

Hunt, D.R. & Taylor, N.P. (eds.). 1986. The genera of the Cactaceae: Towards a new consensus. Bradleya 4: 65-78.

Hunt, D.R. & Taylor, N.P. (eds.). 1990. The genera of Cactaceae: progress towards consensus. Bradleya 8: 85-107.

Johnson, L.A. & Soltis, D.E. 1994. matK DANN sequences and phylogenetic reconstruction in Saxifragaceae s.str. Syst. Bot. 19: 143-156.

Kimnach, M. 1960. A revision of Borzicactus. Cact. Succ. J. (Los Angeles) 32: 8-13, 57-60, 92-96, 19-112.

Maddison, D.R. & Maddison, W.P. 2000. MacClade version 4: Analysis of phylogeny and character evolution. Sunderland, MA, USA: Sinauer Associates.

Meregalli, M., Ercole, E. & Rodda, M. 2010. Molecular phylogeny vs. morphology: shedding light on the infrageneric classification of *Gymnocalycium* (Cactaceae). Biodivers. Ecology 3: 257-27.

Metzing, D. & Kiesling, R. 2008. The study of cactus evolution: the pre-DNA era. Haseltonia 14: 6-25.

Mosti, S., Lewke Bandara, N. & Papini, A. 2011. Further insights and new combinations in *Aylostera* (Cactaceae) based on molecular and morphological data. Pakistan J. Bot. 43: 2769-2785.

Mottram, R. 2006. Osservazioni sulle Borzicactinae. Observations in Borzicactinae. Cactus & Co. 10: 183-130.

Nyffeler, R. 2002. Phylogenetic relationships in the cactus family (Cactaceae) based on evidence from trnK/matK and trnL-trnF sequences. Amer. J. Bot. 89: 312-326.

Nyffeler, R. & Eggli, U. 2010. A farewell to dated ideas and concepts - molecular phylogenetics and a revised suprageneric clssification of the family Cactaceae. Biodivers. Ecol. 3: 109-149.

Oxelman, B., Liden, M. & Berglund, D. 1997. Chloroplast *rps16* intron phylogeny of the tribe Sileneae (Caryophyllaceae). Pl. Syst. Evol. 206: 393-410.

Prager, E.M. & Wilson, A.C. 1988. Ancient origin of lactalbumin from lysozyme: analysis of DNA und amino acid sequences. J. Mol. Evol. 27: 326-335.

Ritter, F. 1980. Kakteen in Südamerika. Vol 2. Argentinien / Bolivien. Spangenberg, D: Published by the author.

Ritter, F. 1981. Kakteen in Südamerika. Vol. 4. Peru. Spangenberg, D: Published by the author.

Ritz, C.M., Martins, L., Mecklenburg, R., Germykin, V. & Hellwig, F.H. 2007. The molecular phylogeny of *Rebutia* (Cactaceae) and its allies demonstrates the influence of paleogeography on the evolution of South American mountain cacti. Amer. J. Bot. 94: 1321-1332.

Schlumpberger, B.O. 2012. New combinations in the *Echinopsis* alliance. Cact. Syst. Init. 28: 29-31.

Schlumpberger, B.O. & Renner, S.S. 2012. Molecular phylogenetics of *Echinopsis* (Cactaceae): polyphyly at all levels and convergent evolution of pollination modes and growth forms. Amer. J. Botany 99: 1335-1349.

Shimodaira, H. & Hasegawa, M. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16: 1114-1116.

Stamatakis, A. 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22: 2688-2690.

Swofford, D.L. 2002. PAUP* Phylogenetic analysis using parsimony (*and other methods), version 4.0b10. Sunderland, MA, USA: Sinauer Associates.

Templeton, A.R. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. Evolution 37: 221-244.

Wallace, R.S. 1995. Molecular systematic study of the Cactaceae: using chloroplast DNA variation to elucidate cactus phylogeny. Bradleya 13: 1-12.

Wallace, R.S. & Gibson, A.C. 2002. Evolution and systematics. In: Nobel, P.S. (ed.), Cacti. Biology and Uses, pp. 1-21. Berkley, CA, US: University of California Press.

Wiens, J.J., Fetzner, J.W., Parkinson, C.L. & Reeder, T.W. 2005. Hylid frog phylogeny and sampling strategies for speciose clades. Syst. Biol. 54: 719-748.

Wiens, J.J. 2006. Missing data and the design of phylogenetic analyses. J. Biomed. Infor. 39: 34-42.

Figures and Tables

Figure 1. Maximum likelihood topology with support values derived from 1'000 rapid bootstrap searches.

Figure 2. Strict consensus tree of 4061 maximum parsimony topologies with our suprageneric classification into monophyletic tribes, subtribes and subclades of Cereeae s.l. mapped onto it.

Table 1. Suprageneric and generic classification of ingroup exemplars of the "BCT" clade based on Hunt & al. (2006) used in this study in comparison to our systematization derived from the phylogenetic analyses and statistical tests.

Table 2. Species sampled for the present study with information on tribal and generic classification by Hunt & al. (2006), voucher information, and GenBank accession codes.

Table 3. Description of molecular markers and aligned matrices concerning the number of ingroup and outgroup exemplars, number of characters in the alignment, number of informative characters, and number of clades resolved in the ingroup of the strict consensus tree. Figures in brackets refer to the number of characters for the ingroup taxa only.

Table 4. Comparing the combined data set with the *trnS-trnG* data set concerning bootstrap support values for some major clades of Cereeae s.l.

Table 5. Parsimony- and likelihood-based paired-sites tests comparing constrained tree topologies representing alternative monophyletic groups tested for 16 taxonomic concepts, using the Templeton test (i.e., Wilcoxon signed-ranks test) and the winning-sites test in a maximum parsimony framework and the Shimodaira-Hasegawa test with full optimization in a maximum likelihood framework. Taxon circumscription is not considered when monophyly is rejected as $P < 0.01$, while monophyly is not rejected as $P > 0.05$.

Figure 1.
Maximum likelihood topology

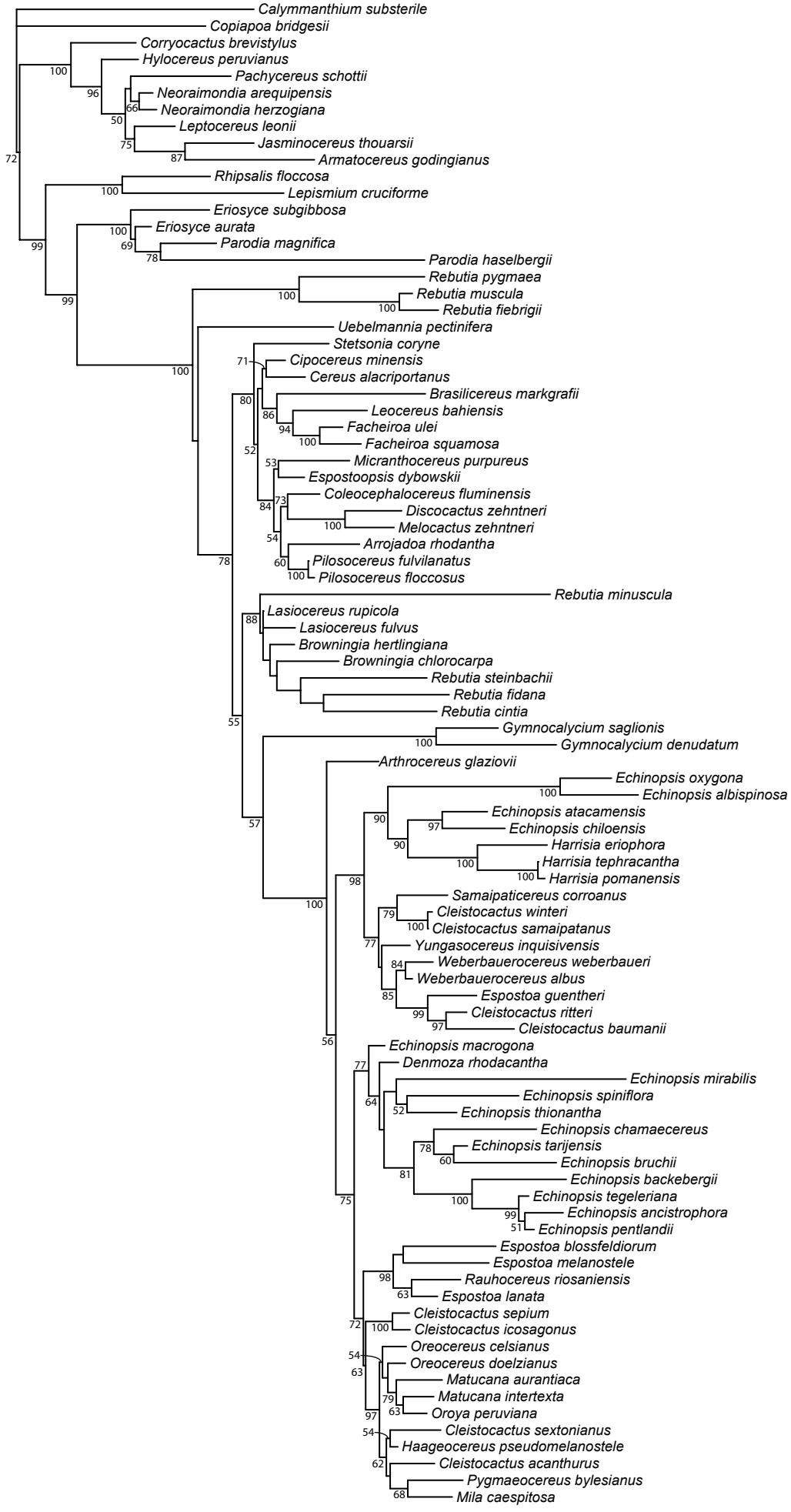


Figure 2.
Strict consensus tree

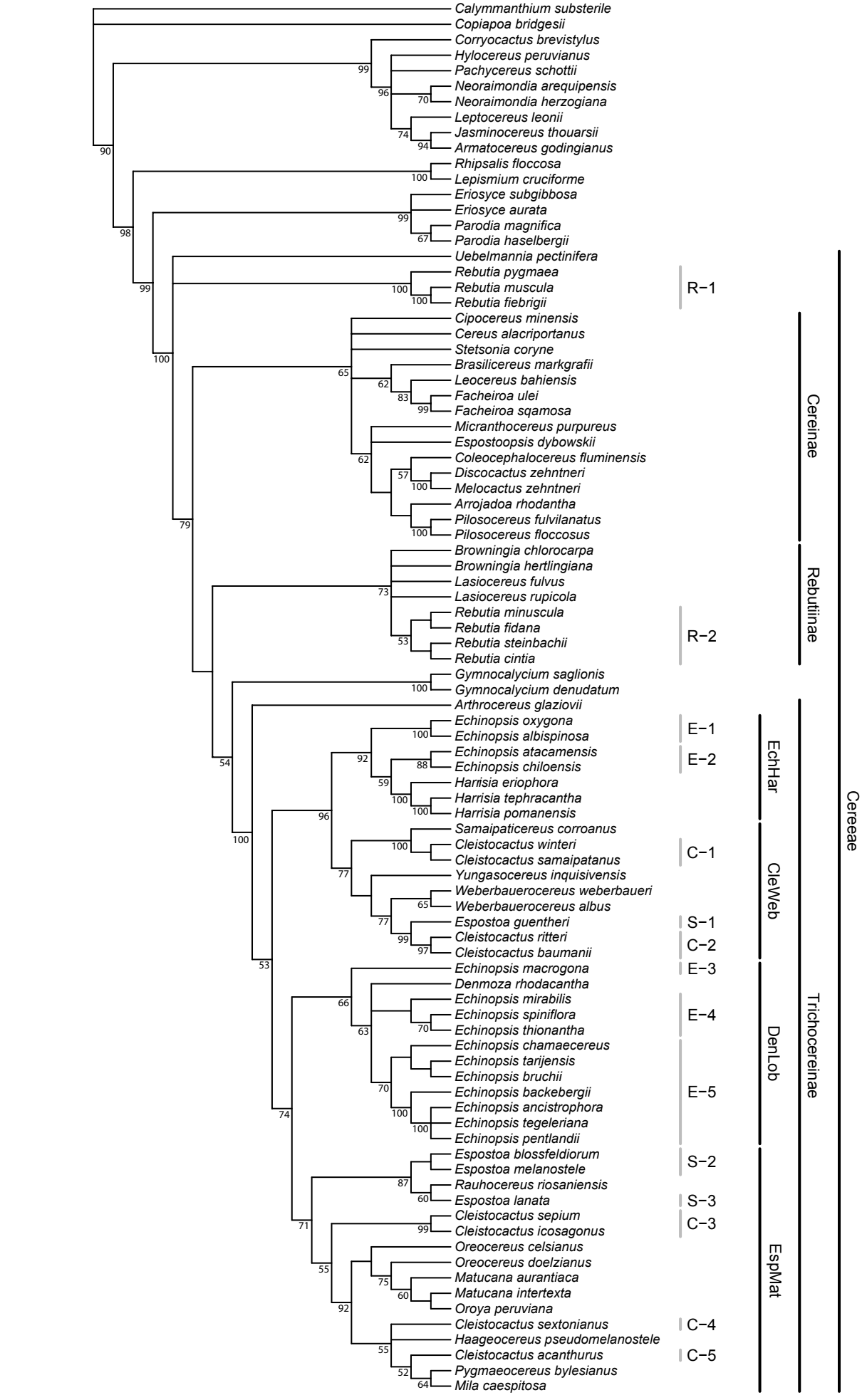


Table1. Suprageneric and generic classification

Hunt et al. 2006		Schlumpberger & Renner 2012	our classification (Lendel et al., submitted)	
suprageneric taxa ingroup taxa (BCT)	generic classification	Schlumpberger 2012, Hunt 2012	suprageneric taxa	generic classification
Cereeae	Arrojadoa rhodantha		Cereinae	Arrojadoa rhodantha
Cereeae	Brasilicereus markgrafii		Cereinae	Brasilicereus markgrafii
Cereeae	Browningia chlorocarpa		Browningiinae	Browningia chlorocarpa
Cereeae	Browningia hertlingiana		Browningiinae	Browningia hertlingiana
Cereeae	Cereus alacriportanus		Cereinae	Cereus alacriportanus
Cereeae	Cipocereus minensis		Cereinae	Cipocereus minensis
Cereeae	Coleocephalocereus fluminensis		Cereinae	Coleocephalocereus fluminensis
Cereeae	Espostoopsis dybowskii		Cereinae	Espostoopsis dybowskii
Cereeae	Melocactus zehntneri		Cereinae	Melocactus zehntneri
Cereeae	Micranthocereus purpureus		Cereinae	Micranthocereus purpureus
Cereeae	Pilosocereus floccosus		Cereinae	Pilosocereus floccosus
Cereeae	Pilosocereus fulvilanatus		Cereinae	Pilosocereus fulvilanatus
Cereeae	Stetsonia coryne		Cereinae	Stetsonia coryne
Cereeae	Uebelmannia pectinifera		incertae sedis	Uebelmannia pectinifera
Trichocereeae	Arthrocereus glaziovii	Arthrocereus glaziovii	Trichocerinae	Arthrocereus glaziovii
Trichocereeae	Cleistocactus acanthurus	Loxanthocereus acanthurus	Trichocerinae (EspMat)	Loxanthocereus acanthurus
Trichocereeae	Cleistocactus baumanii	Cleistocactus baumanii	Trichocerinae (CleWeb)	Cleistocactus baumanii
Trichocereeae	Cleistocactus icosagonus	? Borzicactus icosagonus	Trichocerinae (EspMat)	Borzicactus icosagonus
Trichocereeae	Cleistocactus ritteri	Cleistocactus ritteri	Trichocerinae (CleWeb)	Cleistocactus ritteri
Trichocereeae	Cleistocactus samaipatanus	?	Trichocerinae (CleWeb)	Bolivicereus samaipatanus
Trichocereeae	Cleistocactus sepium	Borzicactus sepium	Trichocerinae (EspMat)	Borzicactus sepium
Trichocereeae	Cleistocactus sextonianus	Borzicactus sextonianus	Trichocerinae (EspMat)	Loxanthocereus sextonianus
Trichocereeae	Cleistocactus winteri	?	Trichocerinae (CleWeb)	Bolivicereus winteri
Trichocereeae	Denmoza rhodacantha	Denmoza rhodacantha	Trichocerinae (DenLob)	Denmoza rhodacantha
Trichocereeae	Discocactus zehntneri		Cereinae	Discocactus zehntneri
Trichocereeae	Echinopsis albispinosa	Echinopsis albispinosa	Trichocerinae (EchHar)	Echinopsis albispinosa
Trichocereeae	Echinopsis ancistrophora	Lobivia ancistrophora	Trichocerinae (DenLob)	Lobivia acistrophora
Trichocereeae	Echinopsis atacamensis	Leucostele atacamensis	Trichocerinae (EchHar)	Echinopsis atacamensis
Trichocereeae	Echinopsis backebergii	Lobivia backebergii	Trichocerinae (DenLob)	Lobivia backebergii
Trichocereeae	Echinopsis bruchii	Soehrensia bruchii	Trichocerinae (DenLob)	Lobivia bruchii
Trichocereeae	Echinopsis chamaecereus	Chamaecereus silvestrii	Trichocerinae (DenLob)	Lobivia chamaecereus
Trichocereeae	Echinopsis chiloensis	Leucostele chiloensis	Trichocerinae (EchHar)	Echinopsis chiloensis
Trichocereeae	Echinopsis macrogona	Trichocereus macrogonus	Trichocerinae (DenLob)	Trichocereus macrogonus
Trichocereeae	Echinopsis mirabilis	Setiechinopsis mirabilis	Trichocerinae (DenLob)	Setiechinopsis mirabilis
Trichocereeae	Echinopsis oxygona	Echinopsis oxygona	Trichocerinae (EchHar)	Echinopsis oxygona

Trichocereeae	Echinopsis pentlandii	Lobivia pentlandii	Trichocerinae (DenLob)	Lobivia pentlandii
Trichocereeae	Echinopsis spiniflora	Acanthocalycium spiniflorum	Trichocerinae (DenLob)	Acanthocalycium spiniflorum
Trichocereeae	Echinopsis tarijensis	Soehrensia tarijensis	Trichocerinae (DenLob)	Lobivia tarijensis
Trichocereeae	Echinopsis tegeleriana	Lobivia tegeleriana	Trichocerinae (DenLob)	Lobivia tegeleriana
Trichocereeae	Echinopsis thionantha	Acanthocalycium thionathum	Trichocerinae (DenLob)	Acanthocalycium thionanthum
Trichocereeae	Espostoa blossfeldiorum	Espostoa blossfeldiorum	Trichocerinae (EspMat)	Espostoa blossfeldiorum
Trichocereeae	Espostoa guentheri	Vatricania guentheri	Trichocerinae (CleWeb)	Vatricania guentheri
Trichocereeae	Espostoa lanata	Espostoa lanata	Trichocerinae (EspMat)	Espostoa lanata
Trichocereeae	Espostoa melanostele	Espostoa melanostele	Trichocerinae (EspMat)	Espostoa melanostele
Trichocereeae	Facheiroa squamosa		Cereinae	Facheiroa squamosa
Trichocereeae	Facheiroa ulei		Cereinae	Facheiroa ulei
Trichocereeae	Gymnocalycium denudatum		insertae sedis	Gymnocalycium denudatum
Trichocereeae	Gymnocalycium saglionis		insertae sedis	Gymnocalycium saglionis
Trichocereeae	Haageocereus pseudomelanostele	Haageocereus pseudomelanostele	Trichocerinae (EspMat)	Haageocereus pseudomelanostele
Trichocereeae	Harrisia eriophora	Harrisia taetra	Trichocerinae (EchHar)	Harrisia eriophora
Trichocereeae	Harrisia pomanensis	Harrisia pomanensis	Trichocerinae (EchHar)	Harrisia pomanensis
Trichocereeae	Harrisia tephracantha	Harrisia tetracantha	Trichocerinae (EchHar)	Harrisia tephracantha
Trichocereeae	Lasiocereus fulvus		Browningiinae	Lasiocereus fulvus
Trichocereeae	Lasiocereus rupicola		Browningiinae	Lasiocereus rupicola
Trichocereeae	Leocereus bahniensis		Cereinae	Leocereus bahiensis
Trichocereeae	Matucana aurantiaca	Matucana aurantiaca	Trichocerinae (EspMat)	Matucana aurantiaca
Trichocereeae	Matucana intertexta	Matucana intertexta	Trichocerinae (EspMat)	Matucana intertexta
Trichocereeae	Mila caespitosa	Mila caespitosa	Trichocerinae (EspMat)	Mila caespitosa
Trichocereeae	Oreocereus celsianus	Oreocereus celsianus	Trichocerinae (EspMat)	Oreocereus celsianus
Trichocereeae	Oreocereus doelzianus	Oreocereus doelzianus	Trichocerinae (EspMat)	Oreocereus doelzianus
Trichocereeae	Oroya peruviana	Oroya peruviana	Trichocerinae (EspMat)	Oroya peruviana
Trichocereeae	Pygmaeocereus bylesianus	Pygmaeocereus bylesianus	Trichocerinae (EspMat)	Pygmaeocereus bylesianus
Trichocereeae	Rauhocereus riosaniensis	Rauhocereus riosaniensis	Trichocerinae (EspMat)	Rauhocereus riosaniensis
Trichocereeae	Rebutia cintia		Browningiinae	Weingartia cintia
Trichocereeae	Rebutia fidana		Browningiinae	Weingartia fidana
Trichocereeae	Rebutia fiebrigii		incertae sedis	Aylostera fiebrigii
Trichocereeae	Rebutia minuscula		Browningiinae	Rebutia minuscula
Trichocereeae	Rebutia muscula		incertae sedis	Aylostera muscula
Trichocereeae	Rebutia pygmaea		incertae sedis	Mediolobivia pygmaea
Trichocereeae	Rebutia steinbachii		Browningiinae	Weingartia steinbachii
Trichocereeae	Samaipaticereus corroanus	Samaipaticereus corroanus	Trichocerinae (CleWeb)	Samaipaticereus corroanus
Trichocereeae	Weberbauerocereus albus	Weberbauerocereus albus	Trichocerinae (CleWeb)	Weberbauerocereus albus
Trichocereeae	Weberbauerocereus weberbaueri	Weberbauerocereus weberbaueri	Trichocerinae (CleWeb)	Weberbauerocereus weberbaueri
Trichocereeae	Yungasocereus inquisivensis	Yungasocereus inquisivensis	Trichocerinae (CleWeb)	Yungasocereus inquisivensis

*species combination not yet available

Table 2. Species sampled for the present study

Hunt et al. 2006 suprageneric taxa	Hunt et al. 2006 species classification	specimen voucher	GenBank ID: maturase K (matK) gene	GenBank ID: *rpl16*	GenBank ID: *rps16*	GenBank ID: *trnS-trnG*
outgroup taxa						
Echinocereeae	Calymmanthium substerile	Orig. hort., cult. ZSS 89 3442 /0 (ZSS 32051)	AY015291	submitted	submitted	no seq
Notocacteae	Copiapoa bridgesii	Chile, Chañaral, Knize 1399, cult. ZSS (ZSS 19863)	AY015293	no seq	submitted	no seq
Echinocereeae	Corryocactus brevistylus	Chile, Prov. Iquique, 5 km S of Mamiña, Eggli 2748a (ZSS 18145)	AY015302	submitted	submitted	no seq
Echinocereeae	Armatocereus godingianus	Ecuador, Chimborazo Prov., Alausi - Huigra, Supthut 89103 (ZSS 28228)	AY015296	no seq	submitted	no seq
Echinocereeae	Jasminocereus thouarsii	Ecuador, Galápagos, Isabela, Villamil, Supthut 89146 (ZSS 3589)	JX683856	no seq	no seq	no seq
Echinocereeae	Neoraimondia herzogiana	Orig. hort., cult. Botanical Garden Zürich (ZSS s.n.)	AY015315	no seq	submitted	no seq
Echinocereeae	Neoraimondia arequipensis	Peru, Atico, cult. ZSS, Ostolaza 94966, cult. ZSS (ZSS 19861)	AY015299	no seq	submitted	no seq
Echinocereeae	Leptocereus leonii	Cuba, Sierra de Anafe, Areces s.n., cult. ZSS (ZSS 28286)	AY015297	submitted	submitted	no seq
Echinocereeae	Pachycereus schottii	Orig. hort. Mesa Garden (ZSS 19859)	AY015309	submitted	submitted	no seq
Hylocereeae	Hylocereus peruvianus	Peru, near Rioja, 800 m, Rauh 35393, cult. ZSS (ZSS 3628)	AY015310	no seq	submitted	no seq
Rhipsalideae	Lepismium cruciforme	Brazil, Rio Grande do Sul, Caçapava do Sul, Horst & Uebelmann HU 1101, cult. ZSS (ZSS 32048)	AY015344	no seq	submitted	no seq
Rhipsalideae	Rhipsalis floccosa	Argentina, Tucumán, 5 km W of Alpachiri, Leuenberger & Eggli 4643 (ZSS 18976)	AY015342	no seq	submitted	no seq
Notocacteae	Eriosyce subgibbosa	Orig. hort. Mesa Garden (ZSS 19871)	AY015338	submitted	submitted	no seq
Notocacteae	Eriosyce aurata	Orig. hort., cult. Botanical Garden Zürich (ZSS 19925)	AY015336	no seq	submitted	no seq
Notocacteae	Parodia haselbergii	Orig. hort., cult. Botanical Garden Zürich (ZSS 19924)	AY015330	submitted	submitted	no seq
Notocacteae	Parodia magnifica	Orig. hort., cult. Botanical Garden Zürich (ZSS 19873)	AY015332	submitted	submitted	no seq
ingroup taxa (BCT)						
Cereeae	Arrojadoa rhodantha	Brazil, Minas Gerais, E of Mato Verde, Charles GC 517.01, cult. ZSS (ZSS 28378)	JX683842	submitted	submitted	submitted
Cereeae	Brasilicereus markgrafii	Orig. hort., cult. Botanical Garden Zürich (ZSS s.n.)	JX683870	submitted	submitted	submitted
Cereeae	Browningia chlorocarpa	Peru, Lambayeque / Piura, E of Olmos, Ritter 290, cult. ZSS (ZSS 6043)	AY015316	submitted	submitted	submitted
Cereeae	Browningia hertlingiana	Peru, Ayacucho, Huanta, Knize 334, cult. ZSS (ZSS 19869)	AY015315	submitted	submitted	submitted
Cereeae	Cereus alacriportanus	Brazil, Rio Grande do Sul, São Pedro do Sul, Río Toropi, Eggli et al. 2493, cult. ZSS (ZSS 28225)	Y015313	submitted	submitted	submitted
Cereeae	Cipocereus minensis	Brazil, Minas Gerais, Gouvea - Diamantina, Horst & Uebelmann HU 101 (1982), cult. ZSS (ZSS 3039)	JX683867	submitted	submitted	submitted
Cereeae	Coleocephalocereus fluminensis	Brazil, Rio de Janeiro, 4 km SE of Campos, Supthut 8893, cult. ZSS (ZSS 28227)	AY015318	submitted	submitted	submitted
Cereeae	Espostoopsis dybowskii	Brazil, Bahia, 6 - 7 km E of Porto Alegre, Taylor & al. 1551, cult. ZSS (ZSS 27284)	JX683854	submitted	submitted	submitted
Cereeae	Melocactus zehntneri	Brazil, Bahia, Brejinho das Ametistas, Horst & Uebelmann HU 269, cult. ZSS (ZSS 28230)	JX683849	submitted	submitted	submitted
Cereeae	Micranthocereus purpureus	Orig. hort. Uhlig-Kakteen, cult. ZSS (ZSS 28328)	no seq	submitted	submitted	submitted
Cereeae	Pilosocereus floccosus	Orig. hort., cult. Botanical Garden Zürich (ZSS s.n.)	JX683847	submitted	submitted	submitted
Cereeae	Pilosocereus fulvilanatus	Brazil, Minas Gerais, Santa Barbara, Horst & Uebelmann HU 546, cult. Hort. W. Uebelmann (ZSS 28418)	JX683850	submitted	submitted	no seq

Hunt et al. 2006 suprageneric taxa	Hunt et al. 2006 species classification	specimen voucher	GenBank ID: maturase K (matK) gene	*rpl16*	*rps16*	*trnS-trnG*
Cereeae	Stetsonia coryne	Argentina, Catamarca, Paclín, Leuenberger & Eggli 4361, cult. ZSS (ZSS 15832)	AY015320	submitted	submitted	submitted
Cereeae	Uebelmannia pectinifera	Brazil, Minas Gerais, Engenheiro Dolabeloa, Horst & Uebelmann HU 550, cult. ZSS (ZSS 28239)	AY015319	submitted	submitted	submitted
Trichocereeae	Arthrocereus glaziovii	Brazil, Minas Gerais, SW of Belo Horizonte, Horst & Uebelmann HU 330, cult. ZSS (ZSS 24412)	JX683846	submitted	submitted	submitted
Trichocereeae	Cleistocactus acanthurus	Peru, Matucana, Knize 241, cult. ZSS (ZSS 27260)	no seq	no seq	submitted	submitted
Trichocereeae	Cleistocactus baumanii	Argentina, La Rioja, 4 km NW of Chamical, Leuenberger & Eggli 4183b, cult. ZSS (ZSS 28224)	JX683877	submitted	submitted	submitted
Trichocereeae	Cleistocactus icosagonus	Ecuador, Loja, near Loja, Madsen 50276, cult. ZSS 89 1507/0 (ZSS 32082)	JX683866	submitted	submitted	submitted
Trichocereeae	Cleistocactus ritteri	Orig. hort., cult. ZSS 76 2202 /0 (ZSS 27186, 28265)	JX683864	submitted	submitted	submitted
Trichocereeae	Cleistocactus samaipatanus	Orig. hort., cult. ZSS 84 1075 /0 (ZSS 27204, 27285)	JX683873	submitted	submitted	submitted
Trichocereeae	Cleistocactus sepium	Ecuador, Cañar, c. 3 km from Cañar, Madsen 61097, cult. ZSS (ZSS 27286)	JX683852	submitted	submitted	submitted
Trichocereeae	Cleistocactus sextonianus	Peru, Arequipa, between Chaviña and S-wards to Atico, Ritter 317 loc. 1, cult. ZSS (ZSS 27244)	no seq	no seq	submitted	submitted
Trichocereeae	Cleistocactus winteri	Orig. hort., cult. ZSS 80 3277/ 0 (ZSS 27250)	no seq	no seq	submitted	submitted
Trichocereeae	Denmoza rhodacantha	Argentina, Catamarca, 56 km W of Fiambalá, Leuenberger & Eggli 4677 (ZSS 18991)	JX683840	submitted	submitted	submitted
Trichocereeae	Discocactus zehntneri	Brazil, Bahia, W of Morro de Chapéu, Horst & Uebelmann HU 222, cult. ZSS (ZSS 27282)	JX683848	submitted	submitted	submitted
Trichocereeae	Echinopsis albispinosa	Argentina, Salta, 8 km SE of Campo Quijano, Leuenberger & Eggli 4631, cult. ZSS (ZSS 18864, 27198)	JX683860	no seq	submitted	submitted
Trichocereeae	Echinopsis ancistrophora	Argentina, Salta, Campo Quijano, Kiesling s.n., cult. ZSS (ZSS 27190)	no seq	no seq	submitted	submitted
Trichocereeae	Echinopsis atacamensis	Argentina, Catamarca, 12 km above Choya towards Mina Capillitas, Leuenberger & Eggli 4651 (ZSS 18983)	no seq	no seq	submitted	submitted
Trichocereeae	Echinopsis backebergii	Peru, Huancavelica, La Mejorada, Rausch 396, cult. ZSS (ZSS 26817)	no seq	no seq	submitted	submitted
Trichocereeae	Echinopsis bruchii	Orig. hort., cult. A. Lendel (ZSS 32049)	no seq	no seq	submitted	submitted
Trichocereeae	Echinopsis chamaecereus	Orig. hort., cult ZSS 85 1839 /2 (ZSS 28450)	no seq	submitted	submitted	submitted
Trichocereeae	Echinopsis chiloensis	Orig. hort., cult. R. Nyffeler KG17-87 (ZSS 19874)	AY015322	submitted	submitted	submitted
Trichocereeae	Echinopsis macrogona	Org. hort., cult. M. Machado (ZSS 32050)	no seq	submitted	submitted	submitted
Trichocereeae	Echinopsis mirabilis	Argentina, Córdoba, Salina Ambargasta, STO 432, cult. ZSS (ZSS 28312, 28402)	no seq	submitted	submitted	submitted
Trichocereeae	Echinopsis oxygona	Paraguay, Guairá, Colonia Independencia, Esser 14509, cult. ZSS (ZSS 28448)	no seq	submitted	submitted	submitted
Trichocereeae	Echinopsis pentlandii	Bolivia, Cochabamba, Epizana, Knize 1249, cult. R. Nyffeler ex Hort. Mesa Garden (ZSS 19858)	AY015323	submitted	submitted	submitted
Trichocereeae	Echinopsis spiniflora	Argentina, Córdoba, Taninga, Kiesling s.n., cult. ZSS (ZSS 27295)	no seq	no seq	submitted	submitted
Trichocereeae	Echinopsis tarijensis	Bolivia, Tarija, near Iscayachi, Anonymus s.n., cult. ZSS ZSS 82 3653 /0 (ZSS 32083)	no seq	no seq	submitted	submitted
Trichocereeae	Echinopsis tegeleriana	Peru, Lima, near Oyón, Rausch 387, cult. ZSS 93 2403 /0 (ZSS 8391)	no seq	no seq	submitted	submitted
Trichocereeae	Echinopsis thionantha	Argentina, Catamarca, 60 km S of Santa Maria, Supthut 8696, cult. ZSS; ZSS 862032 (ZSS 32084)	AY015325	submitted	submitted	submitted

			GenBank ID			
Hunt et al. 2006 suprageneric taxa	Hunt et al. 2006 species classification	specimen voucher	maturase K (matK) gene	*rpl16*	*rps16*	*trnS-trnG*
Trichocereeae	Espostoa blossfeldiorum	Peru, Amazonas, Balsas, RRP 502, cult. ZSS; ZSS 99 5825/0 (ZSS 32085)	no seq	no seq	submitted	submitted
Trichocereeae	Espostoa guentheri	Bolivia, Santa Cruz, Vallegrande, Kiesling s.n., cult. ZSS (ZSS 27261)	JX683871	submitted	submitted	submitted
Trichocereeae	Espostoa lanata	Ecuador, Loja, Catamayo valley, Gdaniec s.n., cult. ZSS (ZSS 28309)	JX683863	submitted	submitted	submitted
Trichocereeae	Espostoa melanostele	Orig. hort., cult. ZSS 83 2021 /0 (ZSS 27262)	no seq	no seq	submitted	submitted
Trichocereeae	Facheiroa squamosa	Brazil, Bahia, 12.5 km S of Junco, Taylor & al. 1385b, cult. ZSS (ZSS 28240)	JX683865	submitted	submitted	submitted
Trichocereeae	Facheiroa ulei	Orig. hort., cult. Hort. J.-L. Nagel (ZSS 28496)	JX683841	submitted	submitted	submitted
Trichocereeae	Gymnocalycium denudatum	Orig. hort. Mesa Garden (ZSS 19870)	AY015317	submitted	submitted	submitted
Trichocereeae	Gymnocalycium saglionis	Argentina, sine loco, Ritter 21a (as var. tilcarense), cult. ZSS (ZSS 27174)	JX683853	submitted	no seq	submitted
Trichocereeae	Haageocereus pseudomelanostele	Peru, Manchay Cañón, Ostolaza 84243, cult. Hort. Mesa Garden (ZSS 19862)	AY015329	submitted	submitted	submitted
Trichocereeae	Harrisia eriophora	Cuba, Pinar del Río, Guanahacabibes, Anonymus s.n., cult. ZSS (ZSS 28449)	no seq	no seq	no seq	submitted
Trichocereeae	Harrisia pomanensis	Argentina, Córdoba, 8 km ESE of Serrezuela, Leuenberger & Eggli 4710, cult. ZSS (ZSS 18994)	AY015324	submitted	submitted	submitted
Trichocereeae	Harrisia tephracantha	Bolivia, Florida, near Pampa Grande, Rente 12, cult. ZSS (ZSS 27263)	no seq	no seq	no seq	submitted
Trichocereeae	Lasiocereus fulvus	Peru, Amazonas, Balsas, Charles 572.05 (ZSS 27228)	JX683843	submitted	submitted	submitted
Trichocereeae	Lasiocereus rupicola	Peru, Cajamarca, San Marcos, Charles 560.01 (ZSS 27226)	JX683844	submitted	submitted	no seq
Trichocereeae	Leocereus bahniensis	Brazil, Bahia, 1.5 km W of Seabra, Eggli 1287, cult. ZSS (ZSS 28242)	JX683874	submitted	submitted	submitted
Trichocereeae	Matucana aurantiaca	Peru, La Libertad, Currunday, Ritter 164 loc. 2 (as M. currundayensis), cult. ZSS (ZSS 27264)	no seq	no seq	submitted	submitted
Trichocereeae	Matucana intertexta	Peru, Balsas, Río Marañón, Knize 1153, cult. ZSS (ZSS 18764)	AY015327	submitted	submitted	submitted
Trichocereeae	Mila caespitosa	Peru, Lima, Santa Clara, Knize 243, cult. ZSS (ZSS 27302)	JX683872	submitted	submitted	submitted
Trichocereeae	Oreocereus celsianus	Bolivia, Tarija, Cieneguillas, Knize 889, cult. Hort. Mesa Garden (ZSS 19872)	AY015328	submitted	submitted	submitted
Trichocereeae	Oreocereus doelzianus	Peru, Huancavelica, Villa Azul, below Colcabamba, Ritter 1309 loc. 1, cult. ZSS (ZSS 27187)	no seq	no seq	submitted	submitted
Trichocereeae	Oroya peruviana	Orig. hort., cult. ZSS 96 1191 /0 (ZSS 27171)	JX683875	submitted	submitted	submitted
Trichocereeae	Pygmaeocereus bylesianus	Peru, Arequipa, Matarani, Knize 1058, cult. ZSS (ZSS 27297)	JX683862	submitted	submitted	submitted
Trichocereeae	Rauhocereus riosaniensis	Peru, Rio Santa, Knize 1751, cult. Hort. Mesa Garden (ZSS 19860)	AY015326	submitted	submitted	submitted
Trichocereeae	Rebutia cintia	Orig. hort., cult. ZSS 99 2174 /0 (ZSS 27321)	JX683861	submitted	submitted	submitted
Trichocereeae	Rebutia fidana	Orig. hort., cult. Botanical Garden Zürich (ZSS s.n.)	JX683855	submitted	submitted	submitted
Trichocereeae	Rebutia fiebrigii	Bolivia, Tarija, Escayachi, Knize 861 (as R. spinosissima), cult. ZSS (ZSS 28262)	JX683857	submitted	submitted	submitted
Trichocereeae	Rebutia minuscula	B. O. Schlumpberger x108 (M)	no seq	JQ779813	no seq	JQ779515
Trichocereeae	Rebutia muscula	B. O. Schlumpberger x98 (M)	no seq	JQ779819	no seq	JQ779521
Trichocereeae	Rebutia pygmaea	Bolivia, Oruro, near Challapata, Rausch 676a, cult. ZSS 90 3318 (ZSS 32086)	JX683851	submitted	submitted	submitted
Trichocereeae	Rebutia steinbachii	Bolivia, Cochabamba, Chaporé, Kimnach 2736, cult. ZSS (ZSS 28277)	JX683868	submitted	submitted	submitted
Trichocereeae	Samaipaticereus corroanus	Orig. hort., cult. ZSS 90 3741 /0 (ZSS 32087)	AY015321	submitted	submitted	submitted

			GenBank ID			
Hunt et al. 2006 suprageneric taxa	Hunt et al. 2006 species classification	specimen voucher	maturase K (matK) gene	*rpl16*	*rps16*	*trnS-trnG*
Trichocereeae	Weberbauerocereus albus	Peru, Cajamarca, San Marcos, Catagón, Klopfenstein 143 (as W. longicomus), cult. ZSS (ZSS 27280)	JX683869	submitted	submitted	submitted
Trichocereeae	Weberbauerocereus weberbaueri	Orig. hort., cult. Botanical Garden Zürich (ZSS s.n.)	no seq	no seq	submitted	submitted
Trichocereeae	Yungasocereus inquisivensis	Bolivia, La Paz, Nor-Yungas, 7.5 km from Coripata, Kimnach & al. 2597, cult. ZSS (ZSS 27319)	JX683858	no seq	submitted	submitted

Table 3. Description of molecular markers and aligned matrices

molecular marker	*trnK*/*matK*	*rps16*	*rpl16*	*trnS-trnG*	combined cpDNA data set
total no. Terminals	66	84	62	72	90
no. outgroup terminals	16	15	7	0	16
no. ingroup terminals	50	69	55	72	74
no. of characters (aligned)	2534	915	1153	1210	5812
no. of informative characters*	103 (51)	50 (28)	129 (87)	- (296)	578 (462)
no. of constant characters*	2272 (2420)	803 (859)	874 (959)	- (725)	4674 (4963)
no. of ingroup clades resolved in MP strict consensus	22	13	17	62	60

Table 4. Comparing the combined data set with the *trnS-trnG* data set concerning bootstrap support values for some major clades of Cereeae s.l.

Study	Cereinae	Rebutiinae	Trichocereinae	CleWeb	DenLob	EchHar	EspMat
bootstrap support for combined data set	70%	71%	100%	79%	91%	70%	67%
bootstrap support for *trnS-trnG* data set	53%	56%	99%	61%	92%	70%	57%
no. of clades > 50% for combined data set	8	1	36	6	5	11	7
no. of clades > 50% for *trnS-trnG* data set	5	1	31	6	5	8	5
difference in the number [no.] of exemplars between data sets (i.e., no. for combined data set similar or higher than for *trnS-trnG*)	1	1	0	0	0	0	0

Table 5. Parsimony- and likelihood-based paired-sites tests

	Taxonomic concept (test for monophyly)	Publication	Templeton test P value	winning-sites test P value	Shimodaira-Hasegawa test P value	Taxon list T = Table, C = Column
1	tribe Cereeae	Hunt & al. 2006	< 0.0001	< 0.0001	0.000	T1, C1
2	tribe Trichocereeae	Hunt & al. 2006	< 0.0001	< 0.0001	0.000	T1, C1
3	*Cleistocactus* s.l.	Hunt & al. 2006	< 0.0001	< 0.0001	0.000	T1, C2
4	*Echinopsis* s.l.	Hunt & al. 2006	0.0020	0.0017	0.000	T1, C2
5	*Espostoa* s.l.	Hunt & al. 2006	< 0.0001	< 0.0001	0.000	T1, C2
6	*Rebutia* s.l.	Hunt & al. 2006	0.0558	0.0689	0.014	T1, C2
7	*Browningia*		0.7055	0.8506	0.298	*B. chlorocarpa, B. hertlingiana*
8	*Browningia* + *Lasiocereus*		0.6219	0.7423	0.213	*B. chlorocarpa, B. hertlingiana, L. fulvus, L. rupicola*
9	*Borzicactus* s.l.	Kimnach 1960	< 0.0001	< 0.0001	0.000	species of *Bolivicereus, Borzicactus, Loxantho-cereus, Matucana, Oreocereus* as in T1, C5
10	*Borzicactus* s.l., excl. *Bolivicereus*		0.0516	0.0910	0.001	species of *Borzicactus, Loxanthocereus, Matucana,*

						Oreocereus as in T1, C5
11	*Loxanthocereus*	our study	0.8084	1.0000	0.260	*Loxanthocereus* species as in T1, C5
12	*Echinopsis* s.s.	our study	0.5485	0.6900	0.081	*Echinopsis* species as in T1, C5
13	*Espostoa* s.s.	our study	0.5637	0.7011	0.194	*Espostoa* species as in T1, C5, excluding *Rauhocereus*
14	*Haageocereus* s.l.	not adopted in our study	0.8084	1.0000	0.274	including species of *Haageocereus*, *Mila*, and *Pygmaeocereus,* but excluding *Loxanthocereus* as in T1, C5
15	*Matucana* s.l.	not adopted in our study	0.1936	0.3269	0.023	including species of *Matucana* and *Oreocereus*, but excluding *Oroya* as in T1, C5
16	*Matucana*	our study	0.6547	1.0000	0.136	*Matucana* species as in T1, C5 (excluding *Oroya*)

Chapter 2.

Floral evolution in the South American tribe Cereeae s.l. (Cactaceae: Cactoideae): Pollination syndromes in a comparative phylogenetic context

Anita Lendel [1], Reto Nyffeler [1,3] & Urs Eggli [2]

[1] Institute of Systematic Botany, University of Zürich, Zollikerstrasse 107, CH-8008 Zürich, Switzerland
[2] Sukkulenten-Sammlung Zürich, Mythenquai 88, CH-8002 Zürich, Switerland.
[3] author for correspondence

Manuscript for Plant Systematics and Evolution

Abstract

Floral syndromes have been used as a proxy for inferring pollinators of flowers for many decades. For cacti, four major pollinator guilds (melittophily, ornithophily, sphingophily and chiropterophily) have been recognized, and many individual studies have confirmed that the predicted pollen vectors are indeed regular floral visitors. At the same time, studies have also shown that the seemingly specialized pollination systems predicted by the floral syndromes can be much more generalized in some cases. Here, we investigate the evolution of floral syndromes in a phylogenetically well-understood and morphologically diverse clade of cacti, tribe Cereeae s.l.. The tribe Cereeae s.l. of subfamily Cactoideae embraces 46 genera and c. 590 species, and is an almost exclusively South American clade. Some of its taxa are conspicuous vegetation-dominating plants, and others enjoy popularity in horticulture due to their easy cultivation and colorful flowers. Past generic classifications have largely relied on various characters of flowers (color, size, time of anthesis, symmetry, indumentum of the pericarpel), as well as on differences in growth form. Recent phylogenetic studies have shown that previous classifications, whether recognizing few and widely circumscribed or numerous and narrowly defined genera, are artificial. Based on a phylogenetic backbone derived from a previous molecular study, we here explore possible evolutionary scenarios for several floral characters. We are able to show that all the investigated characters (time of anthesis, floral symmetry, presence of a cephalium, position of the flowers along the axis, indumentum of the pericarpel and perianth tube, pollination syndrome, and flower color) show extensive homoplasy, and utilizing any of them for taxonomy is therefore leading to artificial systems. For all the investigated characters, we identify few to numerous repeated switches from one state to another, with little or no directionality, and all of them have consequently to be regarded as evolutionary labile. As far as pollination syndromes are concerned, an early transition from diurnal melittophily to nocturnal chiropterophily is inferred, and a large extent of the backbone of our phylogeny of Cereeae s.l. shows chiropterophily as the prevalent pollination syndrome, with numerous subsequent shifts to diurnal ornithophily, melittophily and backwards to nocturnal sphingophily and chiropterophily. Based on published observational data, we stress that actual pollination systems are not completely congruent with the inferred pollination syndromes, i.e. there is only a partial overlap between the observed pollinator guilds with the pollinator guilds inferred from the pollination syndromes. Several of the investigated species have pollination systems that involve more than one pollinator guild, i.e. their flowers show parallel adaptations to more than one pollinator group. Using pollination syndromes as a proxy for actual pollination systems is therefore prone to misinterpretations. Spatial organization of flowering, i.e. flowers appearing in order or scattered along the stem, is discusses regarding its possible implications on reproductive success, growth form evolution, and pollination systems.

Keywords: Cereeae, Cereinae, Trichocereinae, South America, flower morphology, flower position, pollination ecology, pollination syndromes, phylogeny, ancestral state reconstruction

Introduction

Successful pollination is a key prerequisite for subsequent fertilization of the ovules, and thus for seed production. The study of flower-pollinator interactions enjoys a long history, partly due to the discovery of astounding "key & lock" combinations of flowers and their specific pollinators (including many "textbook" examples to illustrate coevolution between plants and pollinating animals; e.g. Sitte et al. 2002: 770-774, Judd et al. 2008: 122-123), partly due to the inherent importance of sexual reproduction for plant evolution. The concept of "pollination syndromes" developed in the beginning of the 19th century, (for a brief review see e.g., Johnson et Steiner 2000; Ollerton et al. 2009), culminating in the seminal study of Faegri et van der Pijl (1966). "Pollination syndromes" are specific combinations of floral characteristics (including rewards) that are "associated with the attraction and utilization of a specific group of animals as pollinators" (Fenster et al. 2004: 376), and at best simultaneously exclude illegitimate floral visitors, such as pollen thieves or herbivores, that would negatively impact successful reproduction and thus reproductive fitness. The concept of "pollination syndromes" firmly roots in the general assumption that "the flowers of most angiosperms are sufficiently specialized for pollination by particular animal types" (Johnson et Steiner 2000). Hence, the recognition of syndromes results from the convergent and concerted (canalized) evolution of floral traits in these plant taxa. Influential studies such as Grant et Grant (1965) for the phlox family contributed significantly to the general notion that "key & lock" relationships between flowers and their pollinators are generally in existence, and that pollination syndromes have great predictive value as to the vectors involved with the pollination of a given flower. Such a hypothesis is particularly attractive on the background of the general belief amongst evolutionary biologists that natural selection favors specialization (reviewed, e.g., by Waser et al. 1996), causing the misleading interpretation that generalizations are generally rare.

However, a growing body of evidence started to indicate in the past 30 or 40 years that the "key & lock" mechanism between flower types and pollinators is not always as strict as we would like to see it (Fenster et al. 2004), but that the degree of specificity between flower type and pollinators is variable, and thus a matter of degree (e.g., Ollerton et al. 2007: "as with so many things in biology, simple dichotomies (generalized versus specialized) mask a far more complex and variable reality"). There is an obvious incongruence between "assumed" or "predicted" pollinators, and factually observed pollinators, often involving other than the predicted pollinators (Fenster et al. 2004, Ollerton et al. 2009). While "pollination syndromes" are an attractive and neat concept based on the morphological traits of flowers, comparatively easy to observe, the study of "pollination systems" (i.e., the combination of floral traits and the suite of "real" pollinating animal species from one to several functional

groups) is laborious and time-consuming. Data based on firm observations are therefore slow to accumulate, but for cacti, e.g., Ossa et Medel (2011), Schlumpberger et Raguso 2008, Schlumpberger et al. 2009, or Walter (2010) have all provided such data for several species of columnar to globular species from Chile and Argentina, and additional references are found in Table 1. Waser et al. (1996) not only question the generally accepted views of the importance of "key & lock" mechanisms, but also provide an elucidating discussion why generalized pollination systems (i.e. systems that are visited by additional guilds of pollinators than those predicted by relying exclusively on the floral syndrome) would have to be expected on theoretical grounds. The same authors show that most pollination systems are moderately generalized as a rule (but see Fenster et al. 2004 for a conflicting interpretation of the data), and that there is a continuum between generalization and specialization in pollination systems. In this context it should be kept in mind that the fragmentary nature of many pollinator studies (Waser et al. 1996: 1045), concentrating on visitors expected on the base of the floral syndrome (e.g. hummingbirds for red tubular flowers) rather than trying to identify the whole theoretically possible spectrum (e.g. other diurnal organisms such as bees, wasps, or butterflies), tends to over-emphasize specializations in flower-pollinator interactions. As a whole, generalist systems "are complex, and variable in time and space, and not so easily dissected" (Ollerton et al. 2007), and a formal global test of the pollination syndrome hypothesis (Ollerton et al. 2009) has shown that almost no plant species matches defined syndromes completely. Despite these limitations, floral syndromes continue to provide a basis towards understanding mechanisms of floral diversification (Fenster et al. 2004: 375), especially when the syndromes reflect the characters associated with the primary or most important (in respect to pollination quality, rather than numerical abundance, see Fenster et al. 2004) pollinators. The predictive value of pollination syndromes has been found to be strongest for ornithophilous and melittophilous flowers (Danieli-Silva et al. 2012).

Cacti generally have showy, medium-sized to large flowers that show a great diversity in floral traits. Already the early study of Porsch (1938, 1939) recognized that cacti show a wide diversity of pollination syndromes, including flowers seemingly presenting adaptations to diurnal insects (bees [Hymenoptera] and butterflies [Lepidoptera]) as well as nocturnal insects (hawkmoths [Sphingidae, Lepidoptera]), or vertebrates (diurnal hummingbirds and nocturnal nectar-feeding bats). Floral morphology and associated characters such as time of anthesis, rewards and scent, and the pollination syndromes they characterize, are, however, variable even amongst the species of a genus (e.g., *Mammillaria* with generally bee-syndrome flowers, but isolated occurrences of ornithophilous flowers in, e.g., *M. senilis*), and certainly within larger clades. A particularly interesting lineage of cacti concerning floral characters and pollinators is the tribe Cereeae s.l. For this clade all major pollination syndromes known for cacti (e.g. Rowley 1980) are evident. Hence, based on size, color, time of anthesis, and presence/absence of scent, four major syndromes are recognized in recent studies (e.g., Schlumpberger 2012) as well as in this present investigation: brightly colored, diurnal, medium-sized, shortly funnel- or cup-shaped unscented flowers seemingly adapted to bees (melittophily; occasional observations of visits

by butterflies are subsumed under this syndrome); red, diurnal, medium-sized tubular and often zygomorphic unscented flowers seemingly adapted to hummingbirds (ornithophily); pale dull-colored, nocturnal, medium-sized, fleshy, shortly funnel- or cup-shaped foul-smelling flowers seemingly adapted to bats (chiropterophily); white, nocturnal, large, salverform or funnel-shaped perfumed flowers seemingly adapted to hawkmoths (sphingophily). Many individual studies, ranging from casual sightings to detailed observations, have confirmed that all these pollination guilds are involved with cactus flowers (see Table 1). A growing body of evidence is accumulating, however, that "mixed" or "generalized" pollination systems are in existence for several species: Sahley (1996) reports bats as well as hummingbirds as pollinators for *Weberbauerocereus weberbaueri,* with pronounced year-to-year variations, and Aona et al. (2006) report 2 species of sphingids and 1 species of bat as pollinators for *Micranthocereus purpureus.* Alonso-Pedano et Ortega-Baes (2012) found that the sphingophilous flowers of *Echinopsis schickendantzii* are effectively pollinated both by nocturnal moths as well as diurnal bees, and Lara-Rodríguez et al. (2012) tabulate a surprising diversity of hummingbirds servicing Mexican columnar cacti showing the chiropterophilous pollination syndrome. Moreover, pollination syndromes for some species show considerable lability, and are often "incomplete" in their traits usually associated with a certain syndrome (e.g., *Lobivia ancistrophora*, Schlumpberger et Raguso 2008, as *Echinopsis ancistrophora*).

Cereeae s.l., as re-circumscribed by Lendel et al. (in prep.) includes most taxa of the traditional tribes Cereeae and Trichocereeae (e.g., Anderson 2001; Anderson 2005; Hunt et al. 2006), and counts some 46 genera and 590 species (figures from Nyffeler et Eggli 2010). With few exceptions (i.e., *Harrisia*, *Melocactus*, *Pilosocereus*), the clade is restricted to South America, where many of its large-growing columnar species are landscape-dominating elements (e.g., species of *Cereus*, *Echinopsis atacamensis, E. terscheckii*, *Oreocereus celsianus*, *Stetsonia coryne*). Numerous other, smaller growing globose to shortly columnar taxa enjoy pronounced popularity in horticulture and for cultivation amongst hobby cactus collectors. Their horticultural qualities make them "collectibles", and the pronounced interest in the colorful flowers of dwarf species of *Echinopsis* or *Rebutia* (as previously circumscribed by Hunt et al. 2006) have, at an early stage, resulted in fine-grained classification systems both at generic and specific levels based on easily observable floral traits (e.g., Backeberg 1958-1962; Backeberg 1966; see also discussion by Rowley 1980)

With this study we aim to investigate, in a comparative framework, patterns of floral evolution among the major lineages of the tribe Cereeae s.l., with particular emphasis on ancestral state reconstructions of several distinct floral characters. We consider for this project, in addition to the pollination syndrome with its four different states as outlined above, an additional set of five floral traits associated with pollination biology: time of anthesis, floral symmetry, position of the flowers along the stem as well as presence/absence of a cephalium, indumentum on pericarpel and perianth tube, and perianth segment color. We attempt to reconstruct ancestral states for individual floral traits that were traditionally used for classificatory decisions as well as for the recognition of pollination syndromes, based on

a molecular phylogenetic reconstruction, and use this insight to discuss the possible importance of character transitions especially as far as pollination syndromes and other floral traits are concerned.

Material and methods

Taxonomic sampling

We conduct this study on the basis of a dataset of 74 representatives of the tribe Cereeae s.l. in combination with 16 outgroup taxa, resulting in a total of 90 study taxa. All study taxa were coded for five different floral traits as well as for the presence / absence of a cephalium, and the "predominant" pollination syndrome (see below). The selection of terminals was guided by the aims of the study of Lendel et al. (in prep.) that explored alternative taxonomies for tribe Cereeae s.l.. Therefore, an attempt was made to include the type species of all groups that were at one time or another treated at the rank of genus. As a consequence, the range of terminals selected for the phylogenetic study is neither representative for the taxonomic diversity of the clades nor the diversity of floral traits present. If there is a distinct variation in these characters within the respective clades, we comment on this in Table 1.

Phylogenetic analysis

We used MrBayes 3.1.2 (Huelsenbeck et Ronquist 2001; Ronquist et Huelsenbeck 2003) to conduct a Bayesian inference analysis on the combined data set of seven different partitions from the cpDNA markers *trn*K/*mat*K, *rpl*16, *rps*16, and *trn*S-*trn*G (Shaw et al. 2005). For *trn*K/*mat*K, which is widely used to infer relationships among major groups of cacti (e.g., Nyffeler 2002; Edwards et al. 2005), four distinct partitions were differentiated: *trn*K1, first/second codon positions of *mat*K, third codon position of *mat*K, and *trn*K2. The GTR+GAMMA+INV substitution model was set for all seven partitions and parameters were kept unlinked with rate priors set to "variable". All other parameters were let at default settings. Two independent runs, each with four chains, were run for 10 million generations sampling every 100 generations. Convergence was assessed by examining the traces of all parameters and the effective sample sizes in Tracer v1.5 (Rambaut et Drummond 2007). Based on visual examination we decided conservatively to discard the first 2.5 million generations as burn-in.

Ancestral character state reconstructions

We used the Ancstates and Stochchar packages in Mesquite 2.75 (Maddison et Maddison 2011) to reconstruct ancestral character states of selected clades in a likelihood framework. In order to address phylogenetic uncertainty, we randomly selected one thousand topologies from the posterior trees of the Bayesian analysis with the help of a small script using the

statistics package R (R Development Core Team, 2012). The tree topologies and the coded character states (shown in the Table 1) were included in a NEXUS data file and ancestral character reconstructions were run via tracing character states over trees with the unordered one-parameter Mk1 likelihood model applied (Maddison & Maddison, 2012). For the selected nodes in the Bayesian posterior probability topology we compiled from the output of ancestral characters state reconstructions a table providing the following information: percentage of node present, percentage of equivocal and unequivocal reconstructions, state most often reconstructed, and percentage of most often reconstructed state (relative to number of trees with corresponding node). On the one hand we identified a series of consecutive nodes in the backbone of tribe Cereeae as nodes a to d, and on the other hand the lineages of the most recent common ancestors of the subclades Cereinae (1), Rebutiinae (2), Trichocereinae (3), and subclades EchHar (4), CleWeb (5), DenLob (6), and EspMat (7). Furthermore, we compiled character map trees with branches proportionally colored for the different states as well as the share of equivocal reconstructions.

Coding of floral traits and pollination syndromes

All our character state codings are as far as possible based on personal observations of living or herbarium material at the Zürich Succulent Plant Collection (ZSS), or on personal observations of natural populations, and were complemented on the basis of published illustrations and / or descriptions (see comments in Table 1). The following remarks apply:
Perianth segment color: We distinguish between pale versus brightly colored flowers, where "pale" includes white flowers as well as pale greenish, yellowish or pinkish hues. Where a taxon is known to exhibit flower color polymorphism (e.g. *Lobivia ancistrophora*, Schlumpberger et Raguso 2008, as *Echinopsis*), we code for the most often encountered color. Where color frequencies cannot be established, we code for the color that is also exhibited by neighboring terminals in the phylogeny.
Flower symmetry: Floral symmetry (actinomorphic vs. zygomorphic) varies along a grade, and the coding is thus somewhat arbitrary. Zygomorphy has different morphological bases (Buxbaum 1953: 165-168), and can result from the positioning of the flower tube relative to the pericarpel and limb (i.e. curved flower tubes as in *Bolivicereus winteri*), or from the oblique positioning of only the limb (free parts of the perianth elements, e.g. in some *Trichocereus* spp.). To some degree, zygomorphy is also brought about by asymmetrical positioning of the stamens due to gravity (e.g. *Cereus* spp. with laterally positioned flowers). Finally, zygomorphy in some taxa is labile and related to the positioning of the flower on the stem (e.g. *Bolivicereus winteri*, flowers zygomorphic if held at an angle from the vertical, flowers actinomorphic if positioned vertically on the stem; Buxbaum 1974).
Position of the flowers along the stem: Flowers on a cactus stem or branch are either produced in a scattered way, or in an ordered sequence that correlates with the sequence in which the areoles are produced. In taxa with "scattered" flowers, the initiation of flower buds is not related to the age of the areoles that produce them, i.e. flowers can appear from areoles lower down on the stem than areoles that flowered / fruited in a preceding season.

For some of the columnar taxa, evidence for flower position is scant, and published photographs are not necessarily unambiguous. A coding error of "ordered" instead of "scattered" is possible when the illustration by chance shows aggregated flowers only, while the opposite coding error is less likely to occur, but it should be noted that for small globose taxa such as *Lobivia* spp., evaluating this character is difficult.

Presence of a cephalium (see next character) is a special form of "ordered" flower position, insofar as flowers are only produced from "dedicated" areoles that form a special part of the stem / branch. It should be noted, however, that even amongst the cephalium-producing cacti, flower position is variable in the extent as described above, i.e. they can be produced in a scattered way (e.g. *Espostoa* spp., *Facheiroa* spp.), or in "ordered" fashion (e.g. *Discocactus*, *Melocactus*). "Ordered" is independent from the overall position along the length of the stem - "ordered" flowers can (and frequently do) appear close to the stem apex or stem "shoulder" (e.g. *Weingartia fidana, Matucana* spp.), but can also appear from the stem side or at the stem base (e.g. *Aylostera* spp., *Rebutia* spp.).

Presence of a cephalium: When areoles with the ability to produce flowers are limited to certain parts of the cactus body, a so-called cephalium is formed, resulting in a distinct flower-producing zone (Buxbaum 1964). When a cephalium is present, it can be lateral (sometimes referred to as "pseudocephalium") or terminal (sometimes referred to as a "true cephalium") (Buxbaum 1975: 6, footnote). The modifications of the fertile zone can be confined to the areoles proper, which produce different (often weaker) spination or hairs / tufts of wool in comparison with ordinary vegetative areoles (e.g. *Pilosocereus* spp.), or it also affects the morphology (usually in the form of reduction in size and/or increase in number) of the ribs (e.g. *Micranthocereus*, *Melocactus*). Apical cephalia can be periodically interrupted (resulting in ring-like floriferous zones, e.g. *Arrojadoa*), or result from a permanent and irreversible switch from the purely vegetative juvenile phase to the purely generative adult phase (e.g. *Discocactus*, *Melocactus*). For our coding, we only differentiate between lateral and terminal, and do not consider neither the degree of change between vegetative and generative areoles nor the amount to which the rib disposition is involved.

Indumentum of pericarpel and perianth tube: The cactus flower has to be interpreted as a modified shoot, into which the ovary is immersed, and which elongates to form a short to long and conspicuous hypanthium (Barthlott et Hunt 1993) that produces a graded set of perianth elements. As a consequence, the pericarpel (the basal part of the modified stem that envelopes the ovary proper) and the perianth tube are composed of few to very numerous nodes, each with a scale-like to enlarged sepaloid to petaloid leaf, with an areole in the axil. When these areoles stop growth early without producing neither spines nor spine rudiments, the flowers are naked apart from the perianth elements. In many taxa, however, the hypanthial areoles produce spines and/or bristles and/or hairs that more or less cover or completely envelope part or most of the flowers (Buxbaum 1953: 126-127). Coding for presence / absence of this indumentum is somewhat arbitrary as the character varies along a grade, and esp. when only few and/or short hairs are produced (e.g. *Espostoa melanostele*), flowers are prone to become mis-coded as "naked".

Time of anthesis: Apart from general statements in descriptions as to "diurnal" or "nocturnal" flowers, only relatively few more detailed observations are available as to the time of anthesis. It must be emphasized that "time of anthesis" is not a unitary character ("nocturnal" vs. "diurnal") but is composed of two more or less independent characters, i.e. the time of first opening, and the duration; arguably, closing / re-opening of flowers should be treated as yet another independent character. While the time of the first opening of flowers appears to be quite constant for a taxon (pers. obs. from cultivation), the duration can be variable as shown by the independent studies of *Echinopsis chiloensis* by Walter (2010; S end of geographical range, flowers first opening in the evening and remaining open for up to 42 hours depending on temperature) and Ossa et Medel (2011; N end of geographical range; flowers strictly diurnal). Day temperature appears to have a great influence (Walter 2010), but early (vs. late) successful pollination and ovule fertilization could also influence the duration of anthesis, as well as genetic control. We introduce a special state "transitional" to cater for seemingly, nocturnal flowers that extend their anthesis well into the following day.
Pollination syndrome: We follow Rowley (1980) and code the pollination syndromes according to the presence / absence of the following characters:
Melittophily: Flowers diurnal, medium-sized, brightly colored, actinomorphic, broadly funnel- to cup-shaped, unscented. Note that we also code the species of *Rebutia* s.l. as melittophilous, despite the statement by Barthlott et Hunt (1993, no primary references cited) that these species are butterfly-pollinated. We have failed to find direct observational data.
Ornithophily: Flowers diurnal, medium-sized, brightly colored, actinomorphic or zygomorphic, narrowly funnel-shaped to tubular, unscented.
Sphingophily: Flowers nocturnal, large and somewhat "flimsy", pale colored to white, actinomorphic or slightly zygomorphic, broadly to narrowly funnel-shaped or salverform, pleasantly perfumed.
Chiropterophily: Flowers nocturnal, medium-sized to large and firm and fleshy, pale colored, actinomorphic or slightly zygomorphic, broadly funnel-shaped to variously cup-shaped, unpleasantly scented (rotting vegetables, garlic, etc.).
In the absence of published opinions or observational studies, we base our codings on personal observations of living material or published illustrations. Where published data is available, we also consider the observed "real" visitors. We introduce the character state "mixed" for cases where more than one guild of pollinators is recorded. Sometimes, pollination syndromes are incomplete in so far as certain characteristics show misfits (e.g. nocturnal brightly colored flowers or nocturnal white flowers without scent in *Lobivia ancistrophora* [Schlumpberger et Raguso 2008, as *Echinopsis*], or nocturnal nectar production in *Echinopsis atacamensis* with only diurnal pollinators [Viana et al. 2001]).

Results

Phylogenetic analysis

The Bayesian phylogenetic analysis resulted in a majority rule consensus topology of 150'002 posterior trees (Fig. 1). Overall, we yield a very well resolved majority rule consensus tree with most posterior probability support values greater than 90 percent (of a total of 85 clades, 61 clades have support of 90 % or more pp.). The topology was found to be absolutely congruent and very similar to results presented in a separate analysis using maximum parsimony and maximum likelihood analyses (Lendel et al., in prep.). Tribe Cereeae s.l., as well as subtribes Cereinae, Rebutiinae, and Trichocereinae received 100 percent posterior probability support. The genus *Gymnocalycium* is the sister group to Trichocereinae (support 92 % pp.), and this clade is in turn sister to Rebutiinae (support 94% pp.) (Fig. 1).

Ancestral characters state analyses

Ancestral character state reconstructions are illustrated in Figs. 2, 3, and 4, and percentages of reconstructed states for selected nodes are presented in Table 2. In terms of reconstructing the ancestral states of floral characters and pollination syndrome along the backbone of the tribe Cereeae (i.e., ancestral lineage from which extant subclades diverged) we infer the following characteristics: white to pale colored, actinomorphic, flowers most probably with ordered arrangement that are nocturnal and, based on pollination syndrome, primarily visited by bats. In terms of indumentum on pericarpel and tube our reconstruction is uninformative.

Discussion

Perianth segment color (Fig. 2a). We distinguish between pale and brightly colored flowers, because flower color has been used for classificatory decisions in the same manner as time of anthesis (e.g. Backeberg 1958-1962; Backeberg 1966: colored short flowers = *Lobivia*, pale long flowers = *Echinopsis*, colored long flowers = *Pseudolobivia*). Much of the backbone of our phylogeny is resolved as "pale colored" (nodes b, c, d, 1, 2, 3, 4 and 5; see Table 2 for node percentages), followed by occasional back-switches to bright flowers in all clades except 4, and with notable high frequency in clades 6 and 7. The inability of the flower color character to define evolutionary coherent clades is also strongly indicated by the study of Schlumpberger et Raguso (2008) who found color polymorphisms for *Lobivia ancistrophora*. In addition, it should also be borne in mind that white flowers are not necessarily linked with nocturnal pollination syndromes (chiropterophily, sphingophily) but also occur in strictly diurnal insect-pollinated taxa (e.g. *Gymnocalycium* spp., *Acanthocalycium spiniflorum*).

Floral symmetry (Fig. 2b). Actinomorphic flowers are by far the most common flower type throughout the study group, as well as in all outgroup clades employed in our phylogeny, and actinomorphy is unambiguously the ancestral state. Isolated transitions to zygomorphic flowers occur only in clades 5, 6 and 7. Clade 7 is notable as actinomorphy is shown as ancestral state at node 7, but all the more derived nodes are resolved as zygomorphic, with three switches back to actinomorphy. Flower symmetry therefore appears to be a very labile character in these most highly derived lineages of the tribe. Almost all of the taxa with zygomorphy share brightly colored (usually red or orange) flowers, and this combination of characters is invariably linked with ornithophily. Like flower color and time of anthesis, zygomorphy has also been used in the past for classificatory decisions, but the many independent developments of zygomorphic flowers render such a use futile.

Position of the flowers along the stem (Fig. 3a). Whether flowers appear in orderly or "clustered" fashion, or scattered along the stem, has not been extensively investigated in the past, and there were only few attempts to use the character for classificatory purposes. One of the few cases where flower position was used taxonomically (e.g. Backeberg 1966) is the classical circumscriptions of *Sulcorebutia* (with basally clustered flowers) vs. *Weingartia* (with apically clustered flowers). In both cases, flowers appear in orderly sequence, and the taxa that make up the two "genera" form a single clade (identified as *Weingartia*) in our study, corroborating earlier investigations (Ritz et al. 2007; Schlumpberger et Renner 2012). Our state reconstruction resolved “ordered” as ancestral character for the whole tribe Cereeae s.l., as well as for nodes b, c and d, and for subtribe Rebutiinae (Fig. 3a, see Table 2 for node percentages). “Flowers in ordered sequence” are also the ancestral state for the immediate outgroup tribe Notocacteae, while both states occur in the remaining outgroups. Within our study group, there are several independent transitions to “flowers scattered along the stem”, the state that is also resolved as ancestral for subtribe Cereinae (clade1) and the EchHar clade (clade 4); while it was found as the most often reconstructed state for the CleWeb (clade 5) and the *Lobivia* subclade (*L. backebergii* to *L. tegeleriana*) of the DenLob clade (clade 6). Flowers position, i.e. “ordered” vs. “scattered”, at first sight appears to have no special significance, but closer scrutiny of the character reveals the following two possibly significant explanations: “ordered” flowers occur as a massive clusters of simultaneously open flowers, and allow massive floral displays even when the individual flowers are of small size. A comparison of flower position with pollination syndrome / time of anthesis shows a fair amount of correlation – nocturnal flowers are often scattered, while diurnal flowers are often ordered. A further correlation (not statistically investigates in this study) can be identified by casual observation between flower position and flower size, and ordered flowers are usually small. Secondly, the presence of scattered flowers could have an important implication for reproductive output at the whole plant level: cactus areoles (i.e. the spine-producing short-shoots developing in the axils of the vestigial primary leaves along the main axis) usually produce a single solitary flower only (Gibson et Nobel 1986: 113; Barthlott et Hunt 1993: 164). If flowers are developed in strict order sequence, reproductive output in a given year is tightly coupled with the amount of growth (i.e. the number of new areoles

produced long the axis) of a preceding year, and if that preceding year happened to have produced below-average growth, or no growth at all, due to climatic conditions, flowering capacity in a subsequent year is negatively impacted, at least for those taxa with flowers originating from near the apex. Allowing flowers to appear in a scattered way separates flowering capacity from vegetative growth in preceding years, and areoles of lower stem parts that have not yet produced a flower can be activated in a subsequent season - which allows a plant to reproduce even when there should have been no vegetative growth at all in the preceding season(s). This decoupling of vegetative growth in one year, and flowering and thus possible reproductive output in a subsequent year, might have important implications for pollination ecology, as this decoupling presumably also allows a more extended flowering period in comparison with the ordered flowering “flushes”. Further, a scattered flower positioning is also linked with pollinator behavior, and could hypothetically diminish within-plant pollen transfer because the pollinators are forced to travel some distance before encountering the next flower. It is likely that flower position is thus indirectly linked to growth form (esp. size) and also influences the pollinator guilds that could service the flower.

Presence of a cephalium (Fig. 3a). It is obvious from our results that cephalia have arisen independently at least 5 times (*Facheiroa*, clade *Espostoopsis* to *Discocactus*, *Vatricania*, *Espostoa*, *Oreocereus* p.p.) in tribe Cereeae s.l., and with the possible exception of the clade Espostoopsis to Discocactus, the presence of a cephalium cannot be used to diagnostically circumscribe any clade. This corroborates Ritter's view that cephalia have arisen in parallel in different "evolutionary lines" (Ritter 1979: 118). It is therefore evident that *Espostoa* s.l. in the sense of Anderson (2001) or Hunt et al. (2006) is not tenable, nor the even broader concept including *Facheiroa* of Mottram (2006). In the clade *Espostoopsis* to *Discocactus*, it is notable that the cephalium appears to have been secondarily lost in the genus *Pilosocereus*.

Indumentum of pericarpel and perianth tube (Fig. 3b). The ancestral character state in the outgroups as well as the most often reconstructed state for the node Cereeae s.l. is "indumentum present". Within Cereeae s.l., state reconstruction failed to identify an ancestral state for next three nodes (nodes b, c and d), while all the more derived nodes were unambiguously resolved as "indumentum present". It thus appears possible that the indumentum has been lost independently in at least three 3 clades (1, 2 and *Gymnocalycium*), with occasional switches back to "indumentum present" in clades 1 and 2. This falsifies the hypothesis expressed by Buxbaum (1963) that the more highly derived clades should have "reduced" flowers without hairs or spines (and also with a reduced number of internodes) according to his "law of the abbreviation of the vegetative phase" that in essence formulates a direction or "trend" for evolutionary changes. Nyffeler & Eggli (2010) already discussed the difficulties associated with Buxbaum's "law" and found it generally inapplicable.

Anthesis time (Fig. 4a). The timing of floral anthesis is tightly linked with the pollination syndromes, and we find a close match of nocturnal anthesis with chiropterophily and sphingophily. The use of anthesis time for classification (e.g. Backeberg 1958-1962; Backeberg 1966: diurnal flowers = *Lobivia*, nocturnal flowers = *Echinopsis*) is a futile attempt, as the character shows homoplasy throughout the investigated tribe, and taxa such as *Echinopsis chiloensis* with their extended duration well into the next day (Walter 2010; Ossa et Medel 2011) defy a strict classification as nocturnal / diurnal completely.

Pollination syndrome (Fig. 4b). Ever since Alcorn et al. 1961 identified bats, doves and bees as legitimate pollinators of *Carnegiea gigantea*, operating in succession on any one flower, we should be aware that pollination syndromes for cacti do not necessarily reflect the complete spectrum of "real" pollinators. Even though the usefulness of pollination syndromes is questionable as far as a distinction between generalization and specialization is attempted (Waser et al. 1996), and despite the fact that almost no plant conforms completely with the discrete pollination syndromes (Ollerteon & al. 2009), they can still be used to investigate trends in evolutionary transitions or shifts (Fenster et al. 2004).

Our state reconstructions show that the basal-most lineages of tribe Cereeae s.l. (*Uebelmannia*, *Aylostera*, *Mediolobivia*) as well as the immediate outgroup (tribe Notocacteae) show melittophily. In contrast, for much of the backbone of our phylogeny, nocturnal chiropterophily is identified as ancestral state with good support (Fig. 4b, nodes b, c, d, 1 and 2; see Table 2 for node percentages), indicating an early shift from melittophily to chiropterophily for Cereeae s.l. Chiropterophily was also inferred to be ancestral for the North American columnar cacti of tribe Pachycereeae by Fleming et al. (2009: 1031), although without tested phylogenetic backbone. Within the backbone, isolated shifts from chiropterophily to melittophily (*Rebutia, Weingartia*), sphingophily (*Cereus, Discocactus*) or ornithophily (*Arrojadoa, Melocactus*) are apparent in clades 1 and 2, and for the higher-order clades 3 to 6, an assemblage of shifts to sphingophily, melittophily, and ornithophily are found, including possible back-shifts to chiropterophily (e.g. *Samaipaticereus, Yungasocereus* and *Vatricania*) in clade 5. Within clade 7, chiropterophily is again reconstructed as ancestral state for the subclade formed by *Espostoa* and *Rauhocereus*, while for the remaining parts of clade 7, a shift to ornithophily is postulated, with subsequent shifts to melittophily (*Mila*, *Oroya*) or sphingophily (*Pygmaeocereus*) in a small number of taxa. When comparing the occurrence of chiropterophily with overall growth form, it is notable that all taxa with chiropterophilous flowers are medium- to large-size columnar cacti - a phenomenon already noted by Vogel (1990), and easily explainable by the fact that chiropterophilous flowers must be freely presented without near-by obstructing vegetation in order to be accessibly for the bats.

No directionality in the evolution of pollination syndromes amongst and within clades can be inferred from our data, and switches between different syndromes have likely occurred. This contrasts Dobat et Peikert-Holle (1985: 207) who identify only shifts from sphingophily and ornithophily to chiropterophily, but not vice versa, as well as Schlumpberger (2012), who found only 1 shift away from chiropterophily amongst a total of

22 shifts between pollination syndromes for Cactaceae as a whole. The occurrence of *all* four major floral syndromes throughout the study group indicates a high degree of lability at an evolutionary scale, and this is underlined by the considerable number of taxa that code as "transitional" for this character. With this result in hand, it appears not so surprising that several recent studies have found mixed (generalized) pollination systems for many species. Mixed pollination systems should be characterized by suites of floral traits that ensure that the flowers are attractive to all the relevant pollinators. Flowers that are simultaneously adapted to different pollinator guilds (i.e. different functional groups sensu Fenster et al. 2004) necessarily must exhibit floral traits that make them "attractive" to all of the relevant guilds. As a result, this means that such flowers do not completely conform to any of the discrete syndromes defined (found to be the rule rather than the exception by Ollerton & al. 2009). Taxa with such flowers probably have the innate genetic flexibility to evolve one or another of these traits in order to "concentrate" on any one of several functionally different pollinators. In this context, it is probably not all that surprising that chiropterophily was found to be the most likely ancestral pollination mechanism through a large part of the Cereeae s.l. backbone: chiropterophilous flowers are accessible to all four major guilds of pollinators known for cacti, provided that the start and duration of their anthesis allows access to diurnal visitors (birds, bees). *Micranthocereus purpureus* is a good example, showing a primary bat syndrome, but the flowers are also regularly exploited by sphingids (Aona et al. 2006), despite not having any obvious hawkmoth adaptations. Sphingophilous flowers are already somewhat more restrictive, esp. when provided with a long tube, and exclude bats as (regular) accidental visitors because the reward (nectar) is unavailable to them. Ornithophilous and melittophilous flowers are even more restrictive, and likely, usually because of the time of anthesis, completely exclude both bats and sphingids. Before firm conclusions can be drawn, however, it is important to keep in mind that flower visitors, even when occurring regularly, are not necessarily also efficient pollinators, and it could well be that the sphingids that exploit *Micranthocereus purpureus* are merely regular but accidental visitors that should probably even be classified as nectar thieves. On the other hand, the most regularly occurring and/or numerous flower visitors are not necessarily also the most important pollinators (Fenster et al. 2004). Finally, it should be kept in mind, that any adaptation is good enough when it results in organisms that "work reasonably well" - adaptations "never result in a “perfect' organism" but will persist as long as benefits outweigh costs, and as long as benefits on one front are not tied to unbearable draw-backs on another front (Niklas et Spatz 2012: 17).

Floral syndromes can be interpreted as being an adaptation or mechanism to facilitate attraction of and allow access and successful pollination by legitimate flower visitors, and to exclude illegitimate visitors (pollen or nectar thieves, herbivores). Whether a pollination syndrome is "restrictive" and addresses only one guild of pollen vectors, or whether it is "relaxed" and simultaneously addresses different such guilds, should also be seen in the context of year-to-year fluctuations of the animals involved: When a pollinator (or pollinator guild) occurs with reliability at flowering time, and at the same time is a reliable pollen vector, concentrating on this one pollinator (or pollinator guild) could be beneficial, at

least in the short term. When pollinators (or pollinator guilds) are unreliable in occurrence and/or efficiency, concentration on a single pollinator (or pollinator guild) could be detrimental. Valiente-Banuet et al. (1996) and Fleming (2002) both found evidence that the specificity of pollinator-plant interaction is greatest for cactus species in the tropics of North America, and that this specificity decreases with increasing latitude. Valiente-Banuet et al. (1996) and Fleming (2002) argue that the main reason for this finding is the year-to-year predictability of the climate, which is high in the tropics, and diminishes with increasing latitudes, i.e. there are increasing between-year variations in pollinator availability. Data for South American cacti is not yet sufficiently voluminous to allow a similar conclusion but the study of Sahley (1996) of *Weberbauerocereus weberbaueri* certainly points in the same direction: The flowers exhibit a predominant bat-syndrome, and were preferably pollinated by bats in one of the study years, but mostly by hummingbirds in another study year when bats were largely absent due to climatic reasons. The hummingbirds can thus be regarded as representing a "backup" pollination system - they will not contribute to reproductive effectiveness in years with abundant bat visits, but will ensure success in the years when bats are largely or completely absent.

It is notable that published observations for taxa we investigated for our study indicate that many of them appear to have a "backup" pollination system by extending the duration of their primarily nocturnal anthesis well into the following day (e.g. *Echinopsis chiloensis*, Walter 2010; *Weberbauerocereus weberbaueri*, Sahley 1996). Apparent character state mis-matches (e.g. in *Echinopsis atacamensis* with nocturnal nectar production but only diurnal visitors, Viana & al. 2001) could also be interpreted as remnants of ancient multiple-pollinator systems that have become partly disfunctional with time. Such a loss of function could be due to various causes - the pollinator could have gone locally extinct, the pollinator has itself evolved and no longer uses the cactus flower resource, or the cactus taxon has become geographically more widespread and now grows in places where its former pollinator does not occur.

Finally, mis-matches between pollination syndromes and observed "real" pollinators could also have a different operational base, and the question to ask is whether flowers in every case have to be "attractive" for different sets of co-occurring pollinators, or whether it is rather that some pollinators are "sufficiently desperate" to visit "wrong" flowers. This appears to be the case for the North American *Agave palmeri*, whose flowers are bat-adapted and regularly visited by migratory bats that are attracted by the typical rotten-fruit scents. The flowers are, however, also regularly visited by a hawkmoth that primarily pollinates *Datura wrightii* with typical sphingophilous flowers (Riffell et al. 2008). While the hawkmoth has an innate preference for *Datura* flowers, these authors show that olfactory learning allows the animals to exploit the much richer nectar source of the *Agave* flowers. Similar mixed systems could also be in operation for cactus species, but we are not aware of any detail studies of such.

Conclusions and suggestions for future work

We show extensive homoplasy in all floral traits investigated in this study. Floral traits that have been ascribed with "predictive values" and on which taxonomic decisions were based, are found to be evolutionary labile and therefore leading to artificial systems, such as the case of a widely-used system of Backeberg (Backeberg 1958-1962; Backeberg 1966). None of these traditionally used flower characters are suitable for classificatory purposes, and they alone or in combination failed to completely diagnostically circumscribe any of the subclades we found for Cereeae s.l.

Ancestral state reconstruction of pollination modes identifies chiropterophily as likely characteristic for much of the backbone of Cereeae s.l. This is somewhat surprising, as the most basal subclades of Cereeae s.l. (*Uebelmannia*, *Aylostera*, *Mediolobivia*) have diurnal brightly colored flowers and are most likely melittophilous, and bright diurnal melittophilous or ornithophilous flowers are also predominant in the more highly derived clades of Cereeae s.l. Melittophily is also regarded as the ancestral state for our study group by Schlumpberger (2012: 304). Accordingly, an early switch from diurnal insect-pollinated flowers to nocturnal vertebrate-pollinated flowers has to be assumed, with several switches back to diurnal flowers, and in some cases (e.g. Y*ungasocereus, Vatricania, Espostoa*) with secondary reversals back to chiropterophily. This sheds some doubt on the results of Schlumpberger (2012: 307), who found only a single example of a shift away from chiropterophily amongst a total of 22 shifts between pollination syndromes for the whole cactus family. Since pollination syndromes are an incomplete proxy for "real" pollination systems (as exemplified by the many "mixed" syndromes amongst our study taxa), the transitions here and by Schlumpberger (2012) identified as "shifts" should be re-evaluated, and it is possible that at least some of these "shifts" are in fact developments starting from generalized pollination systems by adaptive specialization on one pollination guild out of the two or more guilds originally present.

In the absence of a completely dated phylogeny of Cereeae s.l., it is impossible to infer the timing of evolutionary shifts in pollinators, nor to know whether the shifts have resulted in diversification rate changes that should be associated with such shifts according to DeWitt Smith (2010). It will be a major challenge to disentangle the influence of pollination system shifts from other reasons that influence diversification rates, such as the availability of new spatial niches, escapes to new climatic environments, or climate changes on a global scale, or resulting from geographic reorganizations. Plants with mixed pollination systems are probably less constrained to evolve under such conditions of change.

References

Alcorn SM, McGregor SE, Olin G. 1961. Pollination of Saguaro Cactus by doves, nectar-feeding bats, and honeybees. Science, ser. 2, 133: 1594-1595

Alonso-Pedano M, Ortega-Baes P. 2012. Generalized and complementary pollination system in the Andean cactus *Echinopsis schickendantzii*. Pl Syst Evol 298: 1671-1677

Anderson E. 2001. The cactus family. Timber Press, Portland

Anderson EF (ed. Eggli U). 2005. Das grosse Kakteen-Lexikon. Ulmer, Stuttgart

Aona LYS, Machado M, Pansarin ER, Cardoso de Castro C, Zappi D, Amaral M do CE. 2006. Pollination biology of three Brazilian species of *Micranthocereus* Backeb. (Cereeae, Cactoideae) endemic to the "campos rupestres". Bradleya 24: 39-52

Backeberg C. 1958-1962 Die Cactaceae. Handbuch der Kakteenkunde. 6 vols. Gustav Fischer Verlag, Jena.

Backeberg C. 1966. Das Kakteenlexikon. Enumeratio Diagnostica Cactacearum. Gustav Vischer Verlag, Jena.

Barthlott W, Hunt DR. 1993. Cactaceae. In: Kubitzki K et al. (eds) The families and genera of vascular plants 2. Springer-Verlag, Berlin etc, pp 161-197

Buxbaum F. 1953 Morphology of cacti. Section II. The flower. Abbey Garden Press, Pasadena

Buxbaum F. 1963. Die Kakteenblüte und das "Gesetz der Verkürzung der vegetativen Phase". Kakt and Sukk 14: 2-5, 22-25

Buxbaum F. 1964. Was ist ein Cephalium? Kakt and Sukk 15: 28-31, 43-48

Buxbaum F. 1974. *Loxanthocereus*. *Loxanthocereus aureispinus*. In: Krainz H (ed) Die Kakteen, part 58. Franck'sche Verlagshandlung, Stuttgart

Buxbaum F. 1975. Ergebnisse der Kakteenforschung von 1962 bis 1973. III. Von *Praecereus* zu *Pseudopilocereus* (1). Kakt and Sukk 26: 6-9

Danieli-Silva A, Tesserolli de Souza JM, Donatti AJ, Pamplona Campos R, Vicente-Silva J, Freitas L, Galarda Varassin I. 2012. Do pollination syndromes cause modularity and predict interactions in a pollination network in tropical high-altitude grasslands? Oikos 121: 35-43

DeWitt Smith S. 2010. Using phylogenetics to detect pollinator-mediated floral evolution. New Phytol 188: 354-363

Dobat K, Peikert-Holle T. 1985. Blüten und Fledermjäuse. Bestäubung durch Fledermäuse und Flughunde (Chiropterophilie). Verlag Waldemar Kramer / Senckenbergische Naturforschende Gesellschaft, Frankfurt am Main

Edwards EJ, Nyffeler R, Donoghue MJ. 2005. Basal cactus phylogeny: Implications of *Pereskia* (Cactaceae) paraphyly for the transition to the cactus life form. Amer J Bot 92: 1177-1188

Faegri K, van der Pijl L. 1966. The principles of pollination ecology. Pergamon Press, Oxford

Fenster CB, Armbruster WS, Wilson P, Dudash MR, Thomson JD. 2004. Pollination syndromes and floral specialization. Ann Rev Ecol Evol Syst 35: 375-403

Fleming TH. 2002. Pollination biology of four species of Sonoran Desert cacti. In: Fleming TH, Valiente-Banuet A (eds) Columar cacti and their mutualists. University of Arizona Press, Tucson, pp 207-224

Fleming TH, Geiselman C, Kress J. 2009. The evolution of bat pollination: a phylogenetic perspective. Ann Bot 104: 1017-1043

Gibson AC, Nobel PS. 1986. The cactus primer. Harvard University Press, Cambridge / London

Grant V, Grant KA. 1965. Flower pollination in the *Phlox* family. Columbia University Press, New York

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17: 754-755.

Hunt DR, Taylor N, Charles G. 2006. The new cactus lexicon. Text / Atlas volumes. dh books, Milborne Port

Johnson SD, Steiner KE. 2000. Generalization versus specialization in plant pollination systems. Tree 15: 140-143

Judd WS, Campbell CS, Kellogg EA, Stevens PF, Donoghue MJ. 2008. Plant systematics. A phylogenetic approach. Ed. 3. Sinauer Associates, Sunderland

Lara-Rodríguez NZ, Díaz-Valenzuela R, Martinez-Garcia V, Maurico-Lopéz E, Anaid-Díaz S, Valle OI, Fisher-de-León AD, Lara C, Ortiz-Pulido R. 2012. Redes de interacción colibrí-planta del centro-este de México. Revista Mex Biodivers 83: 569-577

Lendel A, Eggli U, Nyffeler R. (in prep.). Phylogenetics, taxonomy, and evolution of pollination syndromes in the tribe Cereeae s.l., with a particular focus on the subtribe Trichocereinae (Cactaceae – Cactoideae)

Maddison WP, Maddison DR. 2011 Mesquite: A Modular System for Evolutionary Analysis. Version 2.75. http://mesquiteproject.org

Mottram R. 2006 Osservazioni sulle / Observations in Borzicactinae. Cactus & Co. 10: 183-190

Niklas KJ, Spatz H-C. 2012. Plant physics. University of Chicago Press, Chicago / London

Nyffeler R. 2002. Phylogenetic relationships in the cactus family (Cactaceae) based on evidence from trnK/matK and trnL-trnF sequences. Amer J Bot 89: 312-326

Nyffeler R, Eggli U. 2010. A farewell to dated ideas and concepts - molecular phylogenetics and a revised suprageneric classification of the family Cactaceae. Schumannia 6 = Biodivers Ecol 3: 109-149

Ollerton J, Killick A, Lamborn E, Watts S, Whiston M. 2007. Many ways to be a generalist flower. Taxon 56: 717-728

Ollerton J, Alarcón R, Waser, NM, Price, MV, Watts S, Cranmer L, Hingston, A., Peter CI, Rotenberry J. 2009. A global test of the pollination syndrome hypothesis. Ann Bot 103: 1471-1480

Ossa CG, Medel R. 2011. Notes on the floral biology and pollination syndrome of *Echinopsis chiloensis* (Colla) Friedrich & G. D. Rowley (Cactaceae) in a population of semiarid Chile. Gayana Bot 68: 213-219

Porsch O. 1938. Das Bestäubungsleben der Kakteenblüte I. Cactaceae Jahrb Deutsche Kakt Ges 1938(1): 1-80

Porsch O. 1939. Das Bestäubungsleben der Kakteenblüte II. Cactaceae Jahrb Deutsche Kakt Ges 1939(1): 81-142

R Development Core Team. 2012. R: a language and environment for statistical computing. Version 2.15.2. R Foundation for Statistical Computing, Vienna. Available from: URL http://www.R-project.org.

Rambaut, A. and Drummond A. J. 2007. Tracer v1.4, http://beast.bio.ed.ac.uk/Tracer.

Riffell JA, Alarcón R, Abrell L, Davidowitz G, Bronstein JL, Hildebrand JG. 2008. Behavioral consequences of innate preferences and olfactory learning in hawkmoth-flower interactions. Proc Nation Acad Sci USA 105: 3404-3409

Ritter F. 1979. Kakteen in Südamerika. Ergebnisse meiner 20-jährigen Feldforschungen. Band 1. Published by the author, Spangenberg

Ritz CM, Martins L, Mecklenburg R, Goremykin V,Hellwig FH. 2007. The molecular phylogeny of *Rebutia* (Cactaceae) and its allies demonstrates the influence of paleogeography on the evolution of South American mountain cacti. Amer J Bot 94: 1321-1332

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572-1574

Rowley GD. 1980. Pollination syndromes and cactus taxonomy. Cact Succ J Great Brit 42: 95-98.

Sahley CT. 1996. Bat and hummingbird pollination of an autotetraploid cactus, *Weberbauerocereus weberbaueri* (Cactaceae). Amer J Bot 83: 1329-1336

Schlumpberger BO. 2012. A survey on pollination modes in cacti and a potential key innovation. In: Patiny S (ed) Evolution of Plant-Pollinator Relationships. Cambridge University Press, Cambridge, pp 301-318

Schlumpberger BO, Raguso RA. 2008. Geographic variation in floral scent of *Echinopsis ancistrophora* (Cactaceae): Evidence for constraints on hawkmoth attraction. Oikos 117: 801-814. doi: 10.1111/j.2008.0030-1299.16211.x

Schlumpberger BO, Cocucci, AA, Moré M, Sérsic AN, Raguso RA. 2009. Extreme variation in floral characters and its consequences for pollinator attraction among populations of an Andean cactus. Ann Bot (Oxford) ns 103: 1489-1500

Schlumpberger BO, Renner SS. 2012. Molecular phylogenetics of *Echinopsis* (Cactaceae): Popyphyly at all levels and convergent evolution of pollination modes and growth forms. Amer J Bot 99: 1335-1349

Shaw, J., E. B. Lickey, et al. 2005. "The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis." American Journal of Botany 92(1): 142-166.

Sitte P, Weiler EW, Kadereit JW, Bresinsky A, Körner C. 2002. Strasburger. Lehrbuch der Botanik. Ed. 35. Spektrum Akademischer Verlag, Heidelberg / Berlin

Valiente-Banuet A, Arizmendi MC, Rojas-Martínez A, Domínguez-Canseco L. 1996 Ecological relationships between columnar cacti and nectar-feeding bats in Mexico. J Trop Ecol 12: 103-119

Viana ML de, Ortega-Baes P, Saravia M, Badano EI, Schlumpberger BO. 2001. Biología floral y polinizadores de *Trichocereus pasacana* (Cactaceae) en el Parque Nacional Los Cardones, Argentina. Revista Biol Trop 49: 279-285

Vogel S. 1990. Radiación adaptiva del sindrome floral en las familias neotropicales. Bol Acad Nac Ci Córdoba 59: 5-30

Walter HE. 2010. Floral biology of *Echinopsis chiloensis* ssp. *chiloensis* (Cactaceae): Evidence for a mixed pollination syndrome. Flora (Jena) 205: 757-763

Waser NM, Chittka L, Price MV, Williams NM, Ollerton J. 1996. Generalization in pollination systems, and why it matters. Ecology 77: 1043-1060

Figures and Tables

Figure 1. Bayesian majority rule consensus tree, with outlined posterior probability support values. Clades of the tribe Cereea s.l. are delimited on the right side of the tree according to the classification of Lendel et al. (in prep.)

Figure 2. A 90-taxon Bayesian Majority Rule tree showing ancestral reconstruction of a perianth segment color (2a) and floral symmetry (2b). Letters and numbers next to the nodes correspond to the numbering of the clades: Cereinae (1), Rebutiinae (2), Trichocereinae (3), and subclades EchHar (4), CleWeb (5), DenLob (6), and EspMat (7); as outlined in the text, Figure 1 and Table 2.

Figure 3. A 90-taxon Bayesian Majority Rule tree showing ancestral reconstruction of a position of the flowers along the stem (3a) and floral indumentum (3b). Letters and numbers next to the nodes correspond to the numbering of the clades: Cereinae (1), Rebutiinae (2), Trichocereinae (3), and subclades EchHar (4), CleWeb (5), DenLob (6), and EspMat (7); as outlined in the text, Figure 1 and Table 2.

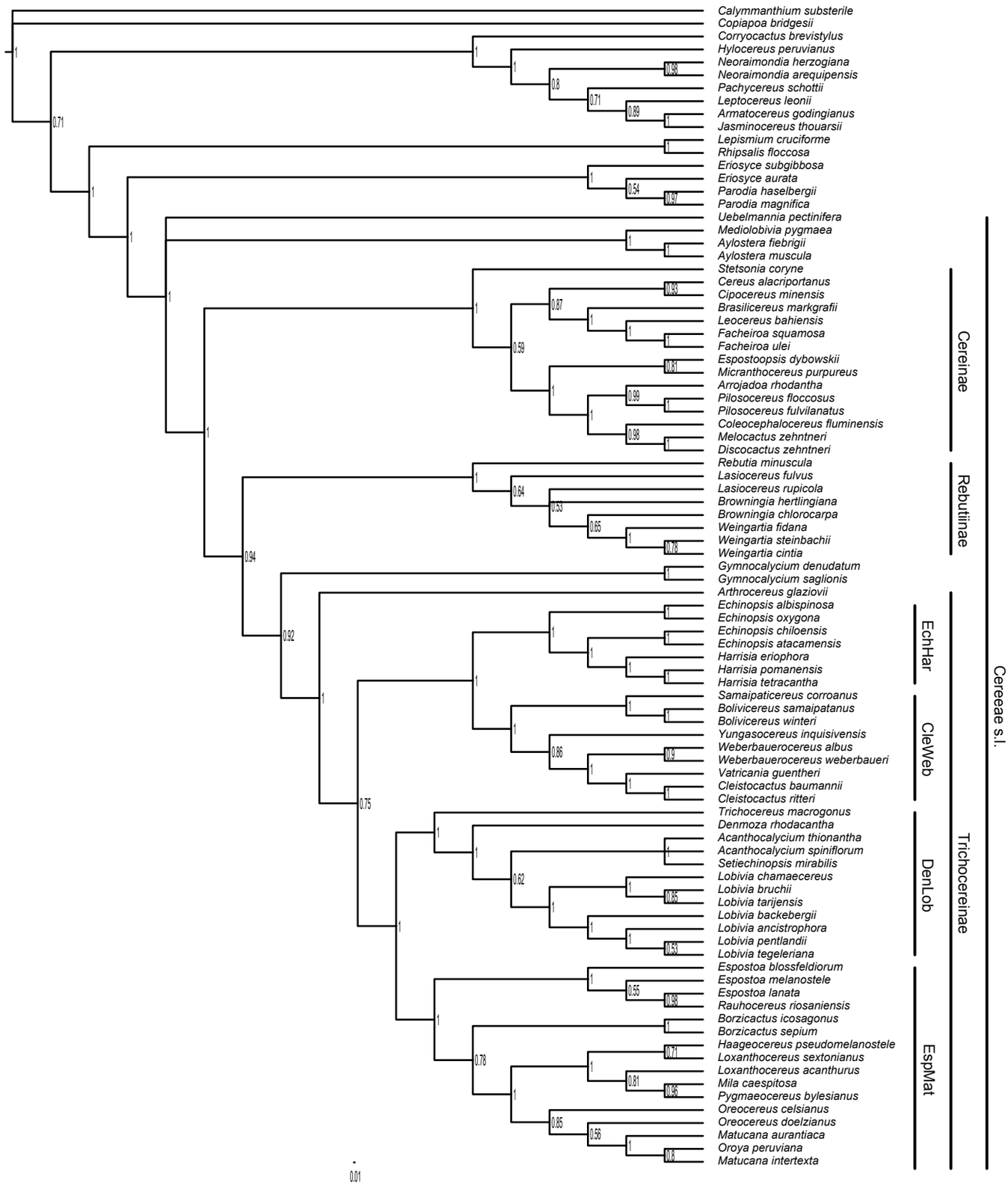
Figure 4. A 90-taxon Bayesian Majority Rule tree showing ancestral reconstruction of the anthesis time (4a) and pollination syndromes (4b). Letters and numbers next to the nodes correspond to the numbering of the clades: Cereinae (1), Rebutiinae (2), Trichocereinae (3), and subclades EchHar (4), CleWeb (5), DenLob (6), and EspMat (7); as outlined in the text, Figure 1 and Table 2.

Table 1. Coding of investigated characters with notes and literature references. Morphological characters in the last column are unordered and coded as follows: color of inner perianth segments (0 = white to pale colored; 1= brightly colored); floral symmetry (0= flowers actinomorphic; 1 = flowers zygomorphic); position of the flowers along the stem (0 = flowers ordered; 1 = flowers scattered along the stem; 2 = flowers arising from a distinct cephalium); cephalium (0 = absent;1 = terminal; 2 = lateral); indumentum on pericarpel and perianth tube (0 = floral tube [and pericarpel] without trichomes; 1 = floral tube with hairs bristles or spines); time of anthesis (0 = flowers primarily nocturnal; 1 = flowers primarily diurnal; 2 = opening time transitional); pollination syndrome (0 = bees; 1 = mixed; 2 = hawkmoths; 3 = hummingbirds; 4 = bats)

Table 2. Node percentages of the ancestral character state reconstructions. Selected nodes are listed according to the naming in the text and Figure 1; and letters and numbers correspond to those in Figs. 2-4. In total 1000 topologies from the posterior trees of the Bayesian analysis were randomly selected and used to reconstruct ancestral character

states of selected clades in a likelihood framework. Cases outlined in gray are the backbone nodes for which a certain state was reconstructed with a very high percentage.

Figure 1. Bayesian majority rule consensus tree



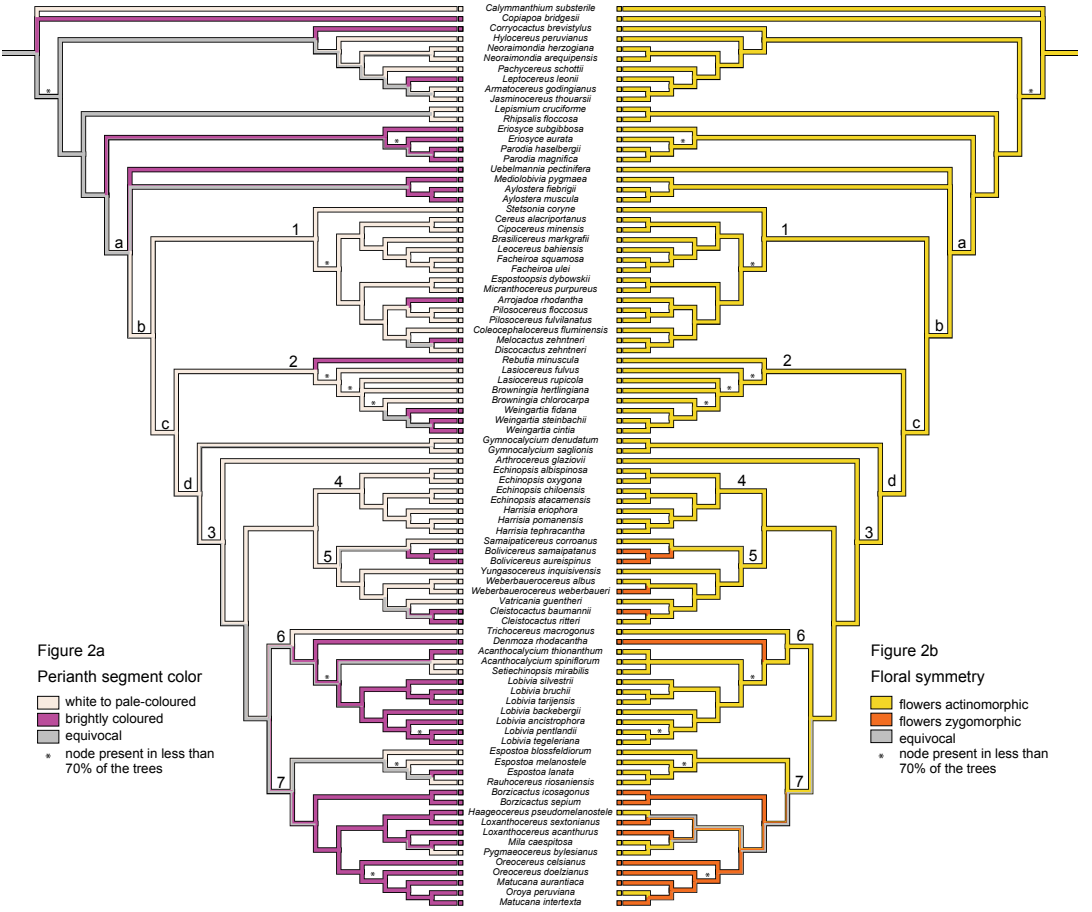


Figure 2a

Perianth segment color

- white to pale-coloured
- brightly coloured
- equivocal
- * node present in less than 70% of the trees

Figure 2b

Floral symmetry

- flowers actinomorphic
- flowers zygomorphic
- equivocal
- * node present in less than 70% of the trees

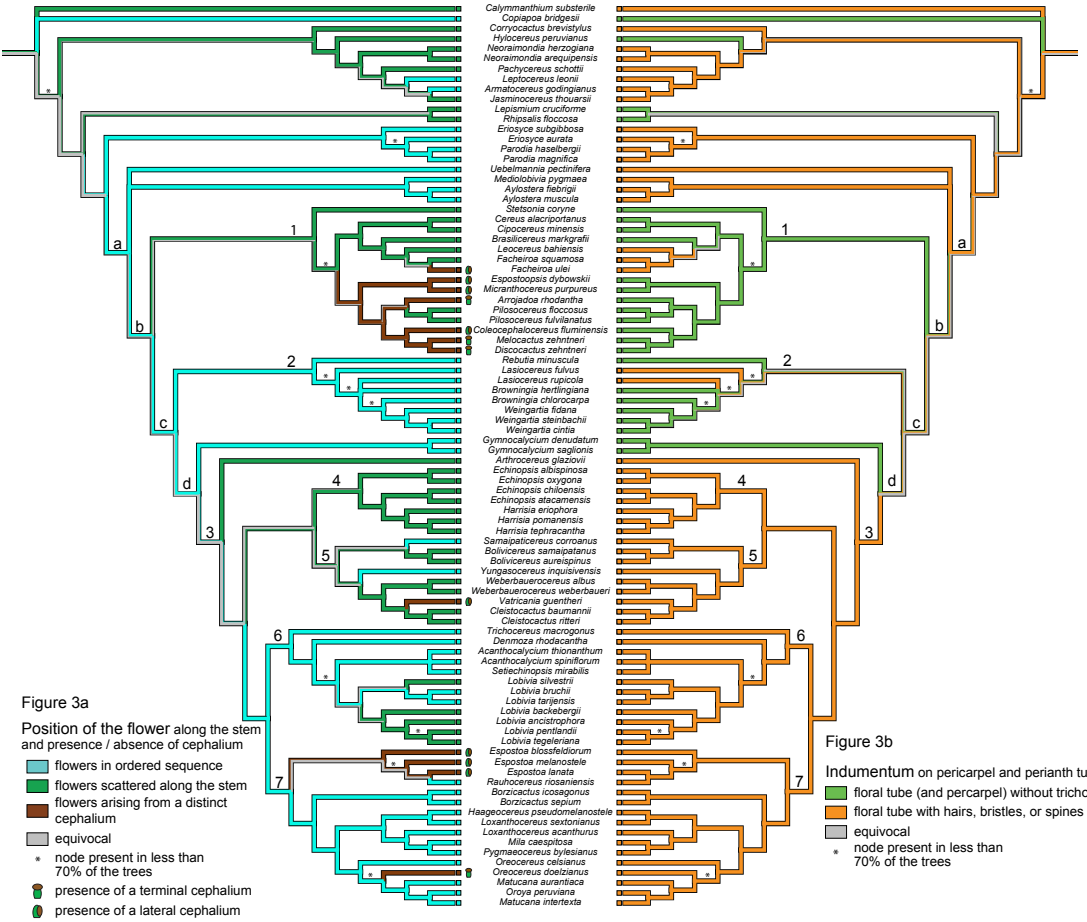


Figure 3a

Position of the flower along the stem and presence / absence of cephalium

- flowers in ordered sequence
- flowers scattered along the stem
- flowers arising from a distinct cephalium
- equivocal
- * node present in less than 70% of the trees
- presence of a terminal cephalium
- presence of a lateral cephalium

Figure 3b

Indumentum on pericarpel and perianth tube

- floral tube (and percarpel) without trichomes
- floral tube with hairs, bristles, or spines
- equivocal
- * node present in less than 70% of the trees

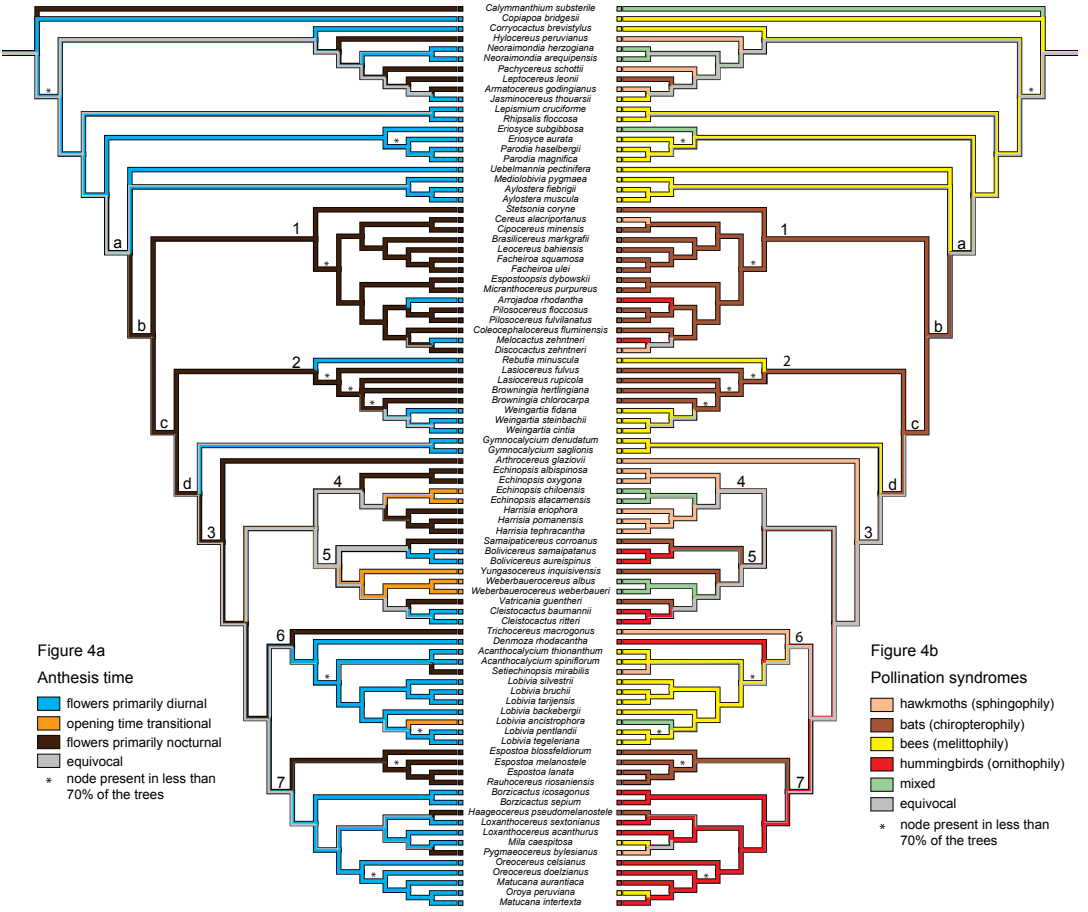


Figure 4a

Anthesis time

- flowers primarily diurnal
- opening time transitional
- flowers primarily nocturnal
- equivocal
- * node present in less than 70% of the trees

Figure 4b

Pollination syndromes

- hawkmoths (sphingophily)
- bats (chiropterophily)
- bees (melittophily)
- hummingbirds (ornithophily)
- mixed
- equivocal
- * node present in less than 70% of the trees

Table 1. Literature references and morphological coding

taxon	inner perianth segment color	position of the flowers along the stemn	indumentum on pericarpel and perianth tube	flower anthesis	pollination syndrome	morphological coding						
Calymmanthium substerile	pers. obs.	pers. obs.	pers. obs., but note that the flower morphology is not directly comparable with that of the other taxa due to the special development of the perianth tube that first forms a closed cap over the flower.	Ritter 1981: 1264; pers. obs.	no data; coded as „mixed“ because the flowers do not completely fit any syndrome.	0	0	1	0	1	0	1
Copiapoa bridgesii	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred from colour and anthesis, while Schulz 2006: 16 casually reports pollination by hover flies.	1	0	0	0	0	1	0
Corryocactus brevistylus	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred from colour and anthesis; Pinto & Kirberg 2009: 100 illustrate a copulating bee on the perianth elements, while Aragón & Aguirre 2007 record bat visits	1	0	1	0	1	1	0
Hylocereus peruvianus	pers. obs.	pers. obs.	inferred from H. monacanthus in Hunt & al. 2006, Atlas fig. 88.1	pers. obs.	inferred from colour and anthesis, and from documented visits of Epiphyllum macropterum with similar flowers (photo L. T. Wasserthal in Eggli 1999: 15), but Valiente-Banuet & al. 1996: 112 report bat visits to flowers of the related Hylocereus undatus.	0	0	1	0	0	0	2
Neoraimondia herzogiana	Hunt & al. 2006, Atlas fig. 15.6	Hunt & al. 2006, Atlas fig. 15.6	Backeberg 1959: 885	inferred	Wood 2007 (Darwin News 8: 3: Dry forest of the Inter-Andean valleys of Bolivia).	0	0	1	0	1	1	1
Neoraimondia arequipensis	Rauh 1958: 260-263, descr.; Ritter 1981: 1269, descr.: colour variable, pale to clear rose-red	Hunt & al. 2006, Atlas fig. 15.4; this taxon shows a special development as areoles produce flowers over many years and develop into a felted spur-shoot	Rauh 1958: 259 for N. roseiflora	Ceroni Stuva & al. 2007	Coded as „mixed“ because the flowers do not completely fit any syndrome, and probably pollinated by a specialized butterfly that also feeds on extrafloral nectaries (Eastwood 2006, Ceroni Stuva & al. 2007)	0	0	1	0	1	1	1
Pachycereus schottii	pers. obs.	pers. obs.	coded for „present“ since this is the condition for Pachycereus s.s., but P. (Lophocereus) schottii differs by having naked flowers (Buxbaum in Krainz 1956-1975: part 44-45 (1970)	pers. obs.	Fleming 1998, mutualistic pollination by the Senita moth (Upiga virescens), related other species (e.g. P. pringlei) with a generalized pollination system involving bats, birds and bees (Fleming & al. 2001)	0	0	1	0	1	0	2
Leptocereus leonii	pers. obs.	pers. obs.	pers. obs.	inferred	inferred from the related L. scopulophilus (Gonzalez-Torres & al. 2011) and from Silva Taboada 1979 for unidentified Leptocereus sp.	1	0	0	0	1	0	4

taxon	inner perianth segment color	position of the flowers along the stemn	indumentum on pericarpel and perianth tube	flower anthesis	pollination syndrome	morphological coding						
Armatocereus godingianus	Rauh 1958: 265, descr.	Hunt & al. 2006, Atlas fig. 12.3	Hunt & al. 2006, Atlas fig. 12.3	Ostolaza 2011: 50, descr., for the genus	inferred by Dobat & Peikert-Holle 1985: 238 for the related A. procerus and A. rauhii, and corroborated from A. cartwrightianus (Arias & al. 2009) and A. procerus (Zamora & al. 2012),	0	0	0	0	1	0	2
Jasminocereus thouarsii	Hunt & al. 2006, Atlas fig. 15.2	inferred from Hunt & al. 2006, Atlas fig. 15.2	Buxbaum in Krainz 1956-1975: part 16, 1961	Jaramillo 2010 (early morning), but see Hunt & al. 2006: 141, described as nocturnal	Jaramillo & al. 2010	0	0	1	0	1	1	0
Lepismium cruciforme	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred from related taxa, e.g. Rhipsalis lumbricoides (Aizen & Feinsinger 1994)	0	0	1	0	0	1	0
Rhipsalis floccosa	pers. obs.	pers. obs.	pers. obs.	pers. obs.	as for Lepismium cruciforme	0	0	1	0	0	1	0
Eriosyce subgibbosa	pers. obs.	pers. obs.	pers. obs.	pers. obs.	humming-birds inferred; bees pers. obs.	1	0	0	0	1	1	1
Eriosyce aurata	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred but note that most species of subgen. Neoporteria are hummingbird-pollinated (Guerrero & al. 2011)	1	0	0	0	1	1	0
Parodia haselbergii	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred	1	0	0	0	1	1	0
Parodia magnifica	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred from the related species P. succinea (Schlindwein & Wittmann 1995).	1	0	0	0	1	1	0
Uebelmannia pectinifera	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred on the base of Schulz & Machado 2000: 30-31, but ornithophily is described by Heek & Strecker 1995.	1	0	0	0	1	1	0
Mediolobivia pygmaea	pers. obs.	pers. obs. - clustered basal	pers. obs.	pers. obs.	Barthlott & Hunt 1993: 167 mention butterfly pollination, but no primary references are cited; butterfly pollination is here subsumed under melittophily; Sahley 1995 infers bees.	1	0	0	0	1	1	0
Aylostera fiebrigii	pers. obs.	pers. obs. - clustered basal	pers. obs.	pers. obs.	as for R. fiebrigii	1	0	0	0	1	1	0
Aylostera muscula	pers. obs.	pers. obs. - clustered basal	pers. obs.	pers. obs.	as for R. fiebrigii	1	0	0	0	1	1	0
Stetsonia coryne	pers. obs.	pers. obs.	pers. obs.	Yetman 2007: 197; pers. obs.	probably bat: Yetman 2007: 197	0	0	1	0	0	0	4
Cereus alacriportanus	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred from related species: hawkmoth for C. alacriportanus (Silva 1995), C. fernambucensis (Locatelli & Machado 1999b), and C. hexagonus (Soriano & Ruiz 2002: 248), but bats for C. repandus (Lemke 1985, sub C. atroviridis: Nassar & al. 1997	0	0	1	0	0	0	2
Cipocereus minensis	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred from the related species C. laniflorus (Ordones Rego & al. 2012), but note that C. pusilliflorus is considered hummingbird-pollinated by Taylor & Zappi 1989: 22.	0	0	1	0	0	0	4

taxon	inner perianth segment color	position of the flowers along the stemn	indumentum on pericarpel and perianth tube	flower anthesis	pollination syndrome	morphological coding						
Brasilicereus markgrafii	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred	0	0	1	0	0	0	4
Leocereus bahiensis	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred by Sahley 1995	0	0	1	0	1	0	4
Facheiroa squamosa	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred by Sahley 1995 for the genus	0	0	1	0	1	0	4
Facheiroa ulei	pers. obs.	pers. obs.	pers. obs.	pers. obs.	as for F. squamosa	0	0	2	2	1	0	4
Espostoopsis dybowskii	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred by Sahley 1995 („bat, moth?“)	0	0	2	2	0	0	4
Micranthocereus purpureus	pers. obs.	pers. obs.	pers. obs.	pers. obs.	Coded as bat syndrome on the base of the overall shape and size of the flower and the primarily nocturnal anthesis, but showing a mixed pollination syndrome: Aona & al. 2006 recorded regular visits by 2 sphingid and 1 bat species, while Conceição & al. 20	0	0	2	2	0	0	4
Arrojadoa rhodantha	pers. obs.	pers. obs.	pers. obs.	pers. obs.	Piedade Kiill & al. 2012; flowers also visited by bees that are considered pollen thieves	1	0	2	1	0	1	3
Pilosocereus floccosus	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred from the related Brazilian species P. catingicola (Locatelli & al. 1997), P. tuberculatus (Rocha 2007), P. aureispinus, P. lanuginosus and P. vilaboensis (Moraes & al. 2005), and from the Venezuelan species P. moritzianus (Nassar & al. 1997)	0	0	1	0	0	0	4
Pilosocereus fulvilanatus	pers. obs.	pers. obs.	pers. obs.	pers. obs.	as for P. floccosus	0	0	1	0	0	0	4
Coleocephalocereus fluminensis	pers. obs.	pers. obs.	pers. obs.	pers. obs.	Porembski 1998 (without primary reference), Ritz & al. 2007: 1325 („probably bat pollinated“)	0	0	2	2	0	0	4
Melocactus zehntneri	pers. obs.	pers. obs.	pers. obs.	pers. obs.	Locatelli & Machado 1999a	1	0	2	1	0	1	3
Discocactus zehntneri	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred	0	0	2	1	0	0	2
Rebutia minuscula	pers. obs.	pers. obs. - clustered basal	pers. obs.	pers. obs.	inferred	1	0	0	0	0	1	0
Lasiocereus fulvus	Ritter 1981: 1479, descr.	Hunt & al. 2006, Atlas figs. 195.1, 195.2	Ritter 1981: fig. 1350	Ritter 1981: 1479	inferred	0	0	0	0	1	0	4
Lasiocereus rupicola	Hunt & al. 2006, Atlas fig. 195.3	Hunt & al. 2006, Atlas figs. 195.3, 195.4	Ritter 1981: fig. 1349	Ritter 1981: 1478	inferred; Schlumpberger 2012: 315 („inferred or own observations“)	0	0	0	0	1	0	4
Browningia hertlingiana	Ritter 1981: 1323, descr.	Anderson 2005: 96	Buxbaum in Krainz 1956-1975: part 31-32, 1965	nocturnal: Ritter 1981: 1322, descr.	inferred, from Aragón & Aguirre 2007 for B. candelaris; but note that Pinto & Kirberg 2009: 234 illustrate a hummingbird visiting the flower of B. candelaris.	0	0	0	0	0	0	4
Browningia chlorocarpa	pers. obs.	Anderson 2005: 96	pers. obs.	pers. obs.	inferred from previous species	0	0	0	0	0	0	4
Weingartia fidana	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred	1	0	0	0	0	1	0
Weingartia steinbachii	pers. obs.	pers. obs. - clustered basal	pers. obs.	pers. obs.	inferred, but see note for Rebutia fiebrigii	1	0	0	0	0	1	0
Weingartia cintia	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred, but see note for Rebutia fiebrigii	1	0	0	0	0	1	0

taxon	inner perianth segment color	position of the flowers along the stemn	indumentum on pericarpel and perianth tube	flower anthesis	pollination syndrome	morphological coding						
Gymnocalycium denudatum	pers. obs.	pers. obs.	pers. obs.	pers. obs.	Schlindwein & Wittmann 1995, Halbritter & al. 1997 argue that all spp. of the genus are bee-pollinated	0	0	0	0	0	1	0
Gymnocalycium saglionis	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred from previous species	0	0	0	0	0	1	0
Arthrocereus glaziovii	Hunt & al. 2006, Atlas fig. 229.1	Hunt & al. 2006, Atlas fig. 229.1	Hunt & al. 2006, Atlas fig. 229.1	inferred	inferredby Jacobi & Fonseca do Carmo 2011, but note that Dobat & Peikert-Holle 1985: 238 infer bats for A. mello-baretoi, and Sahley 1995 infers bats for the genus	0	0	1	0	1	0	2
Echinopsis albispinosa	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred	0	0	1	0	1	0	2
Echinopsis oxygona	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred	0	0	1	0	1	0	2
Echinopsis chiloensis	pers. obs.	pers. obs.	pers. obs.	Walter 2010 (S end of range: nocturnal and diurnal); Ossa & Medel 2011 (N end of range: diurnal)	Walter 2010 (S end of range: sphingids, hummingbirds and bees); Ossa & Medel 2011 (N end of range, several diurnal insects); see also below for E. atacamensis	0	0	1	0	1	2	1
Echinopsis atacamensis	pers. obs.	pers. obs.	pers. obs.	Schlumpberger & Badano 2005	Viana & al. 2001 (bees), Schlumpberger & Badano 2005 (moths, bees, wasps, hummingbird); Ortega-Baez 2011 (moths, bees, hummingbird for the related E. terscheckii)	0	0	1	0	1	2	1
Harrisia eriophora	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred from the related species H. portoricensis (Rojas-Sandoval 2009), but Silva Taboada 1979 mentions bats for an indeterminate Harrisia sp. from Cuba, and González-Oliva & Urquiola 2005 speculate that H. taetra is bat-pollinated	0	0	1	0	1	0	2
Harrisia pomanensis	pers. obs.	pers. obs.	pers. obs.	pers. obs.	as for H. eriophora	0	0	1	0	1	0	2
Harrisia tephracantha	pers. obs.	pers. obs.	pers. obs.	pers. obs.	as for H. eriophora	0	0	1	0	1	0	2
Samaipaticereus corroanus	protologue	Hunt & al. 2006, Atlas fig. 226.1	Hunt & al. 2006, Atlas fig. 226.1	Cárdenas 1952; Ritter 1980: 670	Dobat & Peikert-Holle 1985: 242 (laboratory conditions), Fleming & al. 2009	0	0	0	0	1	0	4
Bolivicereus samaipatanus	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred	1	1	1	0	1	1	3
Bolivicereus aureispinus	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred	1	1	1	0	1	1	3
Yungasocereus inquisivensis	Hunt & al. 2006, Atlas fig. 197.2	Hunt & al. 2006, Atlas fig. 197.2	Hunt & al. 2006, Atlas fig. 197.2	Ritter 1980: 668: open day and night	inferred	0	0	0	0	1	2	4
Weberbauerocereus albus	pers. obs.	Hunt & al. 2006, Atlas fig. 197.4	pers. obs.	pers. obs.; flowers open at dusk and extend anthesis well into the following morning	Sahley 1996	0	0	1	0	1	2	1
Weberbauerocereus weberbaueri	pers. obs.	Hunt & al. 2006, Atlas fig. 199.3	pers. obs.	Sahley 1996	Sahley 1996	0	1	1	0	1	2	1
Vatricania guentheri	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred	0	0	2	2	1	0	4
Cleistocactus baumanii	pers. obs.	pers. obs.	pers. obs.	pers. obs.	Fleming & al. 2009	1	1	1	0	1	1	3
Cleistocactus ritteri	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred	1	0	1	0	1	1	3

taxon	inner perianth segment color	position of the flowers along the stemn	indumentum on pericarpel and perianth tube	flower anthesis	pollination syndrome	morphological coding						
Trichocereus macrogonus	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred, but note that Flores-Saldaña (s.a.) for T. bridgesii infers nocturnal generalist pollination system because nectar production is predominantly nocturnal, but flowers remain open well into the next day	0	0	0	0	1	0	2
Denmoza rhodacantha	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred by Sahley 1995	1	1	0	0	1	1	3
Acanthocalycium thionanthum	pers. obs.	pers. obs.	pers. obs.	pers. obs.	Roig-Alsina & Schlumpberger 2008; pers. comm. M. Giorgetta 2012.	1	0	0	0	1	1	0
Acanthocalycium spiniflorum	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred	0	0	0	0	1	1	0
Setiechinopsis mirabilis	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred	0	0	0	0	1	0	2
Lobivia silvestrii	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred	1	0	1	0	1	1	0
Lobivia bruchii	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred	1	0	0	0	1	1	0
Lobivia tarijensis	pers. obs.	pers. obs.	pers. obs.	pers. obs.	Roig-Alsina & Schlumpberger 2008, Eggli & Giorgetta 2012	1	0	0	0	1	1	0
Lobivia backebergii	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred	1	0	1	0	1	1	0
Lobivia acistrophora	pers. obs.; coded as white as this is the common flower colour in the typical subspecies, but flower colour varies from white to yellow, orange, red and magenta in the species as a whole	pers. obs.	pers. obs.	Schlumpberger & al. 2009: nocturnal to diurnal	Schlumpberger & al. 2009 (some populations mainly sphingids, others solitary bees).	1	0	1	0	1	2	1
Lobivia pentlandii	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred	1	0	1	0	1	1	0
Lobivia tegeleriana	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred	1	0	1	0	1	1	0
Espostoa blossfeldiorum	Ritter 1981: 1485, descr.	Hunt & al. 2006, Atlas fig. 188.2	Buxbaum 1959 (fig. 6); Ostolaza 2011: 131	Ostolaza 2011: 130, for the genus.	inferred	0	0	2	2	1	0	4
Espostoa melanostele	Rauh 1958: 520, descr.; Ostolaza 2011: 134	Hunt & al. 2006, Atlas fig. 191.4	Krainz 1957-1975: part 26, descr. (few very insignificant hairs only)	Rauh 1958: 520	Schlumpberger 2012: 315	0	0	2	2	1	0	4
Espostoa lanata	Rauh 1958: 521, decr.; Ostolaza 2011: 133	Hunt & al. 2006, Atlas fig. 190.3	Rauh 1958: 521, descr.	inferred from related E. melanostele and Ostolaza 2011: 130 for the genus	inferred	1	0	2	2	1	0	4
Rauhocereus riosaniensis	Hunt & al. 2006, Atlas fig. 196.3	Hunt & al. 2006, Atlas fig. 196.3	Hunt & al. 2006, Atlas fig. 196.3	Ostolaza 2011: 236	Dobat & Peikert-Holle 1985: 242 (laboratory conditions), Fleming & al. 2009	0	0	0	0	1	0	4
Borzicactus icosagonus	pers. obs.	Charles 2012	pers. obs.	pers. obs.	inferred	1	1	0	0	1	1	3
Borzicactus sepium	pers. obs.	Hunt & al. 2006, Atlas figs. 212.3, 212.4	pers. obs.	pers. obs.	inferred	1	1	0	0	1	1	3
Haageocereus pseudomelanostele	Hunt & al. 2006, Atlas figs. 186.3, 186.4, flower colour varies from white to carmine-red (ssp. carminiflorus), see also Ritter 1981: 1407, and Calderón & al. 2007	Hunt & al. 2006, Atlas figs. 185.2 – 185.4	Rauh 1958: 427	inferred from H. fascicularis	inferred from Dobat & Peikert-Holle 1985: 240 for the related H. chosicensis and H. horridus	1	0	0	0	1	0	4
Loxanthocereus sextonianus	Hunt & al. 2006, Atlas fig. 213.1	Hunt & al. 2006, Atlas fig. 213.1	Hunt & al. 2006, Atlas fig. 213.1	inferred	inferred	1	1	0	0	1	1	3
Loxanthocereus acanthurus	Hunt & al. 2006, Atlas figs. 208.3, 208.4	Hunt & al. 2006, Atlas figs. 208.3, 208.4	Rauh 1958: 296; Ostolaza 2011: 163	inferred	inferred	1	1	0	0	1	1	3
Mila caespitosa	pers. obs.	Hunt & al. 2006, Atlas figs. 181.1 – 181.3	Rauh 1958: 232-233	pers. obs.	inferred by Sahley 1995	1	0	0	0	1	1	0

taxon	inner perianth segment color	position of the flowers along the stemn	indumentum on pericarpel and perianth tube	flower anthesis	pollination syndrome	morphological coding						
Pygmaeocereus bylesianus	pers. obs.	Hunt & al. 2006, Atlas figs. 182.1, 182.2	Hunt & al. 2006, Atlas fig. 182.1	pers. obs.	inferred	0	0	0	0	1	0	2
Oreocereus celsianus	pers. obs.	pers. obs.	pers. obs.	pers. obs.	Larrea-Alcázar 2011	1	1	0	0	1	1	3
Oreocereus doelzianus	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred	1	1	2	1	1	1	3
Matucana aurantiaca	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred	1	1	0	0	1	1	3
Oroya peruviana	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred by Sahley 1995, although Schlumpberger 2012: 315 casually lists this taxon as being humming-bird pollinated	1	0	0	0	1	1	0
Matucana intertexta	pers. obs.	pers. obs.	pers. obs.	pers. obs.	inferred	1	1	0	0	1	1	3

Table 2. Node percentages of the ancestral character state reconstructions

	node	number of trees that have this node	percentige of equivocal reconstruction	percentige of unequivocal reconstruction	most often reconstructed state	percentage of most often reconstructed state
perianth segment color	a (Cereeae)	1000	99.3	0.7	white to pale colored	0.7
	b	998	0.4	99.6	white to pale colored	99.6
	c	946	0.0	100.0	white to pale colored	100.0
	d	916	0.2	99.8	white to pale colored	99.8
	1 (Cereinae)	997	0.0	100.0	white to pale colored	100.0
	2 (Rebutiinae)	998	0.0	100.0	white to pale colored	100.0
	3 (Trichocereinae)	1000	4.0	96.0	white to pale colored	96.0
	4 (EchHar)	1000	0.0	100.0	white to pale colored	100.0
	5 (CleWeb)	996	0.0	100.0	white to pale colored	100.0
	6 (DenLob)	1000	96.1	3.9	white to pale colored	3.8
	7 (EspMat)	1000	92.1	7.9	brightly colored	7.8
floral symmetry	a (Cereeae)	1000	0.0	100.0	actinomorphic	100.0
	b	998	0.0	100.0	actinomorphic	100.0
	c	946	0.0	100.0	actinomorphic	100.0
	d	916	0.0	100.0	actinomorphic	100.0
	1 (Cereinae)	997	0.0	100.0	actinomorphic	100.0
	2 (Rebutiinae)	998	0.0	100.0	actinomorphic	100.0
	3 (Trichocereinae)	1000	0.0	100.0	actinomorphic	100.0
	4 (EchHar)	1000	0.0	100.0	actinomorphic	100.0
	5 (CleWeb)	996	0.0	100.0	actinomorphic	100.0
	6 (DenLob)	1000	0.0	100.0	actinomorphic	100.0
	7 (EspMat)	1000	33.6	66.4	actinomorphic	66.4
position of the flowers along the stem	a (Cereeae)	1000	4.1	95.9	flowers in ordered sequence	95.7
	b	998	37.1	62.9	flowers in ordered sequence	62.9
	c	946	15.6	84.4	flowers in ordered sequence	84.4
	d	916	30.7	69.3	flowers in ordered sequence	69.3
	1 (Cereinae)	997	16.1	83.9	scattered along the stem	83.9
	2 (Rebutiinae)	998	0.0	100.0	flowers in ordered sequence	100.0
	3 (Trichocereinae)	1000	99.5	0.5	flowers in ordered sequence	0.5
	4 (EchHar)	1000	3.7	96.3	scattered along the stem	96.3
	5 (CleWeb)	996	67.2	32.8	scattered along the stem	32.8
	6 (DenLob)	1000	0.0	100.0	flowers in ordered sequence	100.0
	7 (EspMat)	1000	0.0	100.0	flowers in ordered sequence	100.0
indumentum on pericarpel and perianth tube	a (Cereeae)	1000	24.8	75.2	hairs, bristle or spines	75.1
	b	998	77.2	22.8	hairs, bristle or spines	15.8
	c	946	74.7	25.3	hairs, bristle or spines	18.2
	d	916	78.8	21.2	hairs, bristle or spines	18.7
	1 (Cereinae)	997	0.0	100.0	withouth trichomes	100.0
	2 (Rebutiinae)	998	69.7	30.3	hairs, bristle or spines	19.3
	3 (Trichocereinae)	1000	0.0	100.0	hairs, bristle or spines	100.0
	4 (EchHar)	1000	0.0	100.0	hairs, bristle or spines	100.0
	5 (CleWeb)	996	0.0	100.0	hairs, bristle or spines	100.0
	6 (DenLob)	1000	0.0	100.0	hairs, bristle or spines	100.0
	7 (EspMat)	1000	0.0	100.0	hairs, bristle or spines	100.0
anthesis time	a (Cereeae)	1000	99.5	0.5	nocturnal	0.5
	b	998	1.8	98.2	nocturnal	98.2
	c	946	0.3	99.7	nocturnal	99.7
	d	916	11.4	88.6	nocturnal	88.6
	1 (Cereinae)	997	0.0	100.0	nocturnal	100.0
	2 (Rebutiinae)	998	0.0	100.0	nocturnal	100.0
	3 (Trichocereinae)	1000	41.0	59.0	nocturnal	59.0
	4 (EchHar)	1000	99.3	0.7	nocturnal	0.7
	5 (CleWeb)	996	96.4	3.6	transitional	3.6
	6 (DenLob)	1000	93.2	6.8	nocturnal	6.2
	7 (EspMat)	1000	93.6	6.4	diurnal	4.3
pollination syndrome	a (Cereeae)	1000	99.3	0.7	bats	0.7
	b	998	1.2	98.8	bats	98.8
	c	946	0.0	100.0	bats	100.0
	d	916	37.3	62.7	bats	62.7
	1 (Cereinae)	997	0.0	100.0	bats	100.0
	2 (Rebutiinae)	998	0.0	100.0	bats	100.0
	3 (Trichocereinae)	1000	99.8	0.2	hawkmoths	0.2
	4 (EchHar)	1000	99.9	0.1	hawkmoths	0.1
	5 (CleWeb)	996	90.0	10.0	bats	10.0
	6 (DenLob)	1000	97.5	2.5	hawkmoths	2.2
	7 (EspMat)	1000	65.2	34.8	hummingbirds	34.8

Chapter 3.

Contemporaneous and recent radiations of the world's major succulent plant lineages

Mónica Arakaki[a], Pascal-Antoine Christin[a], Reto Nyffeler[b], **Anita Lendel**[b], Urs Eggli[c], R. Matthew Ogburn[a], Elizabeth Spriggs[a], Michael J. Moore[d], and Erika J. Edwards[a,1]

[a] Department of Ecology and Evolutionary Biology, Brown University, Providence, RI 02912;
[b] Institut für Systematische Botanik, Universität Zürich, CH-8008
Zürich, Switzerland;
[c] Sukkulenten-Sammlung Zürich, Mythenquai 88, CH-8002 Zürich, Switzerland; and
[d] Department of Biology, Oberlin College, Oberlin, OH 44074
[1] To whom correspondence should be addressed. E-mail: erika_edwards@brown.edu

PNAS vol. 108, no 20, 8379 – 8384, 2011

Author contributions: M.A. and E.J.E. designed research; M.A., P.-A.C., R.N., A.L., U.E., R.M.O., E.S., and M.J.M. performed research; M.A., P.-A.C., and E.J.E. analyzed data; and M.A. and E.J.E. wrote the paper.

Abstract

The cacti are one of the most celebrated radiations of succulent plants. There has been much speculation about their age, but progress in dating cactus origins has been hindered by the lack of fossil data for cacti or their close relatives. Using a hybrid phylogenomic approach, we estimated that the cactus lineage diverged from its closest relatives ≈35 million years ago (Ma). However, major diversification events in cacti were more recent, with most species-rich clades originating in the late Miocene, ≈10–5 Ma. Diversification rates of several cactus lineages rival other estimates of extremely rapid speciation in plants. Major cactus radiations were contemporaneous with those of South African ice plants and North American agaves, revealing a simultaneous diversification of several of the world's major succulent plant lineages across multiple continents. This short geological time period also harbored the majority of origins of C_4 photosynthesis and the global rise of C_4 grasslands. A global expansion of arid environments during this time could have provided new ecological opportunity for both succulent and C_4 plant syndromes. Alternatively, recent work has identified a substantial decline in atmospheric CO_2 ≈15–8 Ma, which would have strongly favored C_4 evolution and expansion of C_4-dominated grasslands. Lowered atmospheric CO_2 would also substantially exacerbate plant water stress in marginally arid environments, providing preadapted succulent plants with a sharp advantage in a broader set of ecological conditions and promoting their rapid diversification across the landscape.

Keywords: climate change, paleobotany, CAM photosynthesis

Introduction

Plants are generally classified as succulent when they exhibit pronounced water storage in one or more organs. High degrees of succulence are most often associated with a suite of other characteristics that together confer survival in water-limited environments. This "succulent syndrome" usually includes a shallow root system that permits rapid uptake of unpredictable precipitation; a thick, waxy cuticle that prevents excessive water loss; and Crassulacean acid metabolism (CAM), an alternative photosynthetic pathway that allows plants to uptake atmospheric CO_2 at night when water loss is minimized (Ogburn and Edwards 2010). Although some 30 plant lineages have been classified as succulent, only a small subset of those are species-rich and ecologically important elements of arid and semiarid ecosystems worldwide. These lineages include the ice plants (Aizoaceae, ≈2,000 spp), the spurges (*Euphorbia*, ≈2,100 spp., ≈650 of which are succulent), the stonecrops (Crassulaceae, ≈1,400 spp.), the aloes (*Aloe*, ≈400 spp.), the agaves (*Agave*, ≈200 spp.), the stapeliads and asclepiads (Apocynaceae-Asclepiadoideae, ≈3,700 spp., ≈1,150 of which are succulent) and especially the cacti (Cactaceae, ≈1,850 spp.) (Nyffeler and Eggli 2010b).

The cacti represent the most spectacular New World radiation of succulent plants. Most cacti exhibit a highly specialized life form, with extremely succulent, photosynthetic stems and leaves that have been modified into spines (Gibson and Nobel 1986). The lineage has a broad distribution, but is most prominent in semiarid and arid regions, with several main centers of diversity in arid Mexico and the southwestern United States, the central Andes of Peru and Bolivia, and eastern Brazil (Nyffeler and Eggli 2010b). Despite their ecological importance, the timing of cactus origins and diversification has remained enigmatic. Previous work has emphasized the fact that the cacti are extremely diverse yet almost exclusively New World in distribution, suggesting a possible origin between 90 and 65 Ma, which would allow maximal time for diversification and a spatial separation of Africa and South America (Axelrod 1979, Gibson and Nobel 1986). Others have suggested a more recent origin, because of limited molecular sequence divergence among the major cactus lineages (Hershkovitz and Zimmer 1997, Nyffeler 2002).

There are no relevant fossil records for cacti or their closest relatives, which has made it difficult to estimate divergence times in the group (e.g., Ocampo and Columbus 2010). However, researchers have recently made significant progress in dating the origins of major angiosperm lineages (Moore, et al. 2010), and we exploited these advances to infer the timing of cactus origin and diversification with a two-step approach. First, we sequenced whole chloroplast genomes from 12 cacti and relatives (Table S1) and combined these data with a larger whole-chloroplast data matrix of 90 seed plants (Moore, et al. 2010) to build a broadly sampled phylogeny of angiosperms. We then used multiple fossil calibration points within a Bayesian framework to estimate divergence times and confidence intervals for several key nodes in cacti and relatives (Fig. S1). To look more specifically at patterns and timing of diversification within the major cactus lineages, we performed a series of additional dating analyses on a second phylogeny generated from fewer loci but that included a greatly expanded taxon sampling within the cacti (Fig. 1, Fig. S2, and Table S2). We then identified

the timing of major radiations in cacti and their relatives by implementing a likelihood approach that optimizes the number and placement of shifts in diversification rate across a phylogeny (Alfaro, et al. 2009).

Materials and Methods

Sequence Provenance and Taxon Sampling

A dataset with sequences for 79 protein-coding genes and four ribosomal RNA genes for 90 species of seed plants was obtained from Moore et al. (Moore, et al. 2010). Twelve chloroplast genomes were added to increase sampling in the Caryophyllales (Table S1). Fresh young leaves or photosynthetic stems were obtained from specimens maintained in cultivation at the Brown University Plant Environmental Center or the Sukkulenten-Sammlung Zürich.

To build a phylogeny with better representation of Portulacineae and Cactaceae, plant specimens were obtained from a variety of sources (Table S2). A total of 295 taxa were included. We generated new sequences for 94 taxa for *PHYC* and 63 taxa for *matK/trnK* and combined these with an additional 215 *trnK/matK* plus 22 *PHYC* sequences from the National Center for Biotechnology Information (NCBI). Voucher specimens are deposited at the Brown University Stephen T. Olney Herbarium, the Sukkulenten-Sammlung Herbarium, Zürich; the IADIZA-CRICYT Ruiz Leal Herbarium, Mendoza, or San Marcos University Herbarium, Lima.

Chloroplast Isolation and Sequencing

Chloroplast isolations were performed by using the sucrose gradient centrifugation protocol by Jansen et al. (Jansen, et al. 2005) with a modification for working with succulent plant material. Samples were ground with liquid nitrogen until a coarse powder was obtained, which was quickly transferred to cold Sandbrink isolation buffer. Chloroplast lyses and whole genome amplification were performed with a Qiagen REPLI-g Midi Kit (Qiagen). Library construction and sequencing were performed at the Environmental Genomics Core Facility (EnGenCore) of the University of South Carolina, Columbia. Samples were multiplexed and prepared by following instructions for the 454 GS-FLX instrument (Roche Life Sciences). Raw data (in FASTA format), Newbler preliminary assemblies (from the Newbler software designed for the GS 20 system), quality scores, and NewblerMetrics were received from EnGenCore in standard flowgram format (sff). Alignments and partial assemblies were performed by using MIRA V3rc4 (Chevreux 1997–2010) and Geneious 4.8 (Biomatters). Large contigs or nearly complete assemblies were imported into DOGMA (Wyman, et al. 2004) for annotation, which enabled extraction of individual gene sequences.

DNA Isolation, PCR-Based Amplification, and Sequencing

Total genomic DNA was isolated from fresh or silica gel-dried tissue, using the MP FastDNA SPIN Kit and FastPrep Instrument (MP Biomedicals).

We selected the nuclear phytochrome C (*PHYC*) gene occurring as single copy and shown to be a good source of phylogenetic information (Edwards, et al. 2005). We also incorporated the trnK-maturase K (*trnK/matK*) region, which is the cp region best represented in NCBI and has also proven to be very useful for phylogenetic inference in Cactaceae and other Portulacineae [e.g., Cactaceae (Nyffeler 2002), Montiaceae (O'quinn and Hufford 2005)]. The first exon of *PHYC* (≈1.2 Kb) was amplified by using primers developed by Mathews et al. (Mathews, et al. 2005). The PCR protocol for *PHYC* required a high-quality Taq polymerase (Amplitaq DNA polymerase; Life Technologies) and consisted of a stepdown protocol (with a preheating step of 5 min at 94 °C) beginning at an annealing T of 65 °C and ending at 53 °C, with 2 min annealing, 1 min denaturation at 94 °C, and 1 min primer extension at 72 °C for a total of 36 cycles. Products were cloned by using the StrataClone PCR cloning kit (Agilent Technologies) and sequenced by using the M13 F/R primer pair. Sequencing was performed at the Genomics and Sequencing Center of the University of Rhode Island, using the Applied Biosystem BigDye Terminator v3.1 chemistry. Samples were run on an ABI 3130xl genetic analyzer. Primers and protocol used to amplify and sequence the *trnK/matK* region were developed by Christin et al. (2011b).

Phylogenetic Analyses

The chloroplast dataset of 12 Caryophyllales (Table S1) was added to the broader angiosperm dataset by Moore et al. (Moore, et al. 2010). Nucleotide sequences of protein-coding genes were translated into amino acids, aligned automatically with MUSCLE (Edgar 2004), and adjusted manually with MacClade 4.05 (http://macclade.org). Each gene was aligned separately and later concatenated. Small regions that were difficult to align were excluded from the analysis. The final dataset consisted of 102 taxa, 83 genes, and 75,643 nucleotides. Sequences of *trnK/matK* and *PHYC* were processed as above. Maximum likelihood (ML) analyses of the whole chloroplast matrix and *PHYC*, *trnK/matK*, and combined *PHYC-trnK/matK* datasets were performed with RAxML 7.0.4 (Stamatakis 2006) and run on the Brown University IBM iDataPlex Linux cluster.

Estimation of Divergence Times and Shifts in Diversification

Divergence times were estimated using a two-step approach recommended by Rutschmann (Rutschmann 2006), which applies a Bayesian method that accounts for variable substitution rates between lineages and over time. To start, model parameters were calculated for each of the 83 genes using baseml [PAML package; (Yang 1993)]. Branch lengths and the variance-covariance matrix were then approximated by estbranches (Thorne, et al. 1998). Finally a Bayesian MCMC procedure implemented in multidivtime (multidistribute package; (Thorne, et al. 1998, Thorne and Kishino 2002) was used to estimate posterior distributions of substitution rates and divergence times. The MCMC sampling procedure was run for 1

million generations, after a burn-in of 100,000 generations, with a sampling frequency of 100 generations. We ran two analyses by using 13 fossils as minimum-age node constraints: One analysis used the youngest (lower) bound of the time period to which the fossils were assigned and the second with the oldest (upper) bound. In both cases, constraints were considered minimum ages. We did a sensitivity test (using the youngest assigned ages) that included 13 alternative analyses in which one fossil constraint was excluded at a time. Additional analyses including all fossil constraints but varying MCMC parameters provided similarly consistent results.

Fossils used as calibration points are mostly mesofossils placed with high confidence in their respective phylogenetic position and time frame (Table S3) (Nichols and Traverse 1971, Mapes and Rothwell 1984, Knobloch and Mai 1986, Call and Dilcher 1992, Collinson, et al. 1993, Crane, et al. 1993, Friis, et al. 1994, Hughes 1994, Pedersen, et al. 1994, Crepet and Nixon 1998, Friis, et al. 1998, Sims, et al. 1999, Crepet, et al. 2004, Strömberg 2005, Barreda, et al. 2010, Manchester and O'leary 2010). The fossil record for Caryophyllales is scarce. We considered two candidates: *Coahuilacarpon,* consisting of infructescences of a possible Phytolaccaceae ascribed to the Campanian (Cevallos-Ferriz, et al. 2008), and Amaranthaceae (*Chenopodipollis*) pollen from the early Tertiary (Paleocene) of Texas (Nichols and Traverse 1971). We chose the latter as it is considered to be quite reliably identified (S. Manchester and D. M. Jarzen, personal communication). A well described fossil from the Eocene of Australia has been confidently placed within Caryophyllaceae (Collinson, et al. 1993); however, our taxon sampling did not allow us to make good use of this fossil, because it could only be placed at the same node as *Chenopodipollis*, which is older. Poaceae fossils (e.g., Linder 1987, Crepet and Feldman 1991) are more difficult to place because of insufficient taxon sampling in the present analysis. We assigned a minimum age of 34 Ma to the crown group of Poaceae based on data derived from phytolith analyses (Strömberg 2005). Because this node is likely to be significantly older than that (e.g., Christin, et al. 2008), we also ran a second analysis assigning the same node an age of 65 Ma, as suggested by phytolith morphotypes from dinosaur coprolites (Prasad, et al. 2005). This second analysis did not change our inferences of the age of either Portulacineae or Cactaceae.

Dating analyses for the Cactaceae required a secondary calibration, where age estimates from the analysis of the 83-gene seed plant phylogeny were applied to a well-sampled phylogeny of Portulacineae and outgroups. Upper and lower constraints were set up for Portulacineae and Cactoideae (Fig. S2) by using the estimates obtained from the first two multidivtime analyses, where the oldest and youngest bounds of fossil ages were applied, respectively.

We used a likelihood-based method for identifying shifts in diversification rates by using the program MEDUSA (Alfaro, et al. 2009), as implemented in R. MEDUSA allows users to "fill in" a phylogeny with all extant species by pruning the phylogeny down to the largest possible collection of monophyletic taxa where unsampled taxa may be confidently placed at one of the tips. This approach works well for groups with solid and detailed taxonomic classifications; unfortunately, higher-level taxonomy in the cacti is in a state of

flux, and many of the currently recognized tribes and subtribes are known to be paraphyletic (Nyffeler and Eggli 2010b). Because of this uncertainty, if we were to include every species of Portulacineae in our MEDUSA run, we would be forced to reduce our 295-taxon tree to 42 tips; even worse, the core cacti would have only 5 tips. As an alternative, we used a genus level approach, given that genera circumscriptions have been recently revised and appear to be stabilizing (Anderson 2001, Hunt, et al. 2006, Nyffeler and Eggli 2010b, Nyffeler and Eggli 2010a): We pruned our tree down to one exemplar per sampled genus and then added the total number of species in each genus to the tips (Fig. S3). We did not attempt to include genera that were not present in our 295-taxon tree. Many taxonomically problematic groups were lumped into single large genera to be conservative (e.g., *Echinopsis* includes *Chamaecereus*, *Helianthocereus*, *Lobivia*, *Pseudolobivia*, *Setiechinopsis*, *Soehrensia*, and *Trichocereus*). In the few cases where we had complete sampling for a genus (e.g., *Pereskia*, *Maihuenia*), we included all taxa. This approach enabled us to represent extant diversity quite well, with most major groups attaining >70% coverage and an average coverage across both Cactaceae and Portulacineae of 75% (Table S5). Outside of Portulacineae and Molluginaceae, our sampling was more limited. To ensure that this did not affect inferences inside Cactaceae, we ran MEDUSA only on a tree of Molluginaceae + Portulacineae (Fig. S3). Because of the uncertainty inherent in estimating speciation and extinction rates from phylogenetic topologies, we report MEDUSA rate estimates alongside a statistic assuming a simpler pure-birth model (i.e., assuming zero extinction): D ($D = [\ln(N_t) - \ln(N_o)]/T$ (Table 1 and Table S4), where T is the stem age of the clade, N_t is the number of taxa, and $N_o = 1$; (Baldwin and Sanderson 1998).

Results

Our analysis of 102 chloroplast genomes produced a topology and set of age estimates for major angiosperm nodes that are highly congruent with those of previous studies (Moore, et al. 2007, Moore, et al. 2010) (Fig. S1). Age estimates for particular nodes were extremely robust to removing various combinations of fossil constraints (Table S3). Our phylogenomic analyses suggest that the cacti are ≈35 million years old (Ma), which is much younger than many previous assumptions (Axelrod 1979, Gibson and Nobel 1986) but consistent with speculation based on limited divergence of molecular sequences (Hershkovitz and Zimmer 1997, Nyffeler 2002).

Furthermore, divergence time estimates from our densely sampled phylogeny of cacti and close relatives indicates that many of the important species radiations within this group are actually quite recent (Fig. 1, Table 1, and Table S4). The paraphyletic *Pereskia* comprises the first two diverging, species-poor cactus lineages and are woody trees and shrubs with slightly succulent leaves. The succulent cactus “life-form” emerged in a step-like fashion during early cactus evolution, and certain elements, such as moderate tissue water storage, conservative water use, and variants of CAM photosynthesis, are found in *Pereskia* and other members of the Portulacineae (Edwards and Donoghue 2006, Nyffeler, et al.

2008). Pronounced morphological succulence, as exhibited by the core cacti (Edwards, et al. 2005), did not evolve until ≈25 Ma and was associated with a significant increase in diversification rate (Fig. 1, number 2, and Table 1). However, the most dramatic species radiations in the cacti occurred many millions of years after the evolution of a fully succulent syndrome and were not associated with any obvious anatomical or physiological innovations. We identified five additional shifts in diversification rate in the cacti, the majority occurring within the last 8 Ma (Table 1). With the exception of the genus *Opuntia* (the prickly pears), these shifts occurred at nodes nested within or just outside named taxonomic groups (Nyffeler and Eggli 2010b). Our results suggest that the cactus floras of the three main centers of cactus diversity and endemism (Mexico, central Andes, and Brazil) are extremely young, and more or less contemporary. For example, the North American barrel and columnar cacti both experienced upward shifts in speciation rate roughly 8–6 Ma, which coincides with a similar shift in a clade comprising the Trichocereinae, a South American lineage that comprises the majority of cactus diversity in the central Andean region, and the Cereinae, a nearly exclusively Brazilian lineage (7.5–6.5 Ma; Table 1).

Other noteworthy succulent lineages in our analyses, although located in very different geographical regions, are also of similar age. The endemic Didiereaceae of Madagascar (Didiereaceae *s.s.*) are often called the "Cacti of the Old World," and are stem-succulent trees and shrubs of the spiny-thicket forests. While not being especially species rich (≈12 spp), their diversification began ≈17 Ma; crown *Alluaudia*, the largest and arguably most succulent lineage, is ≈11 Ma. Our crown age estimate of core Ruschioideae (the ice plants, Aizoaceae), an extremely species-rich and fundamental component of the Succulent Karoo flora of South Africa, is ≈17 Ma (Figs. 1 and 2 and Fig. S3). This node age is substantially older than a previous report that suggested a rapid radiation in this group 8.7–3.8 Ma (Klak, et al. 2004); however, our taxon sampling of core Ruschioideae was too sparse to allow for investigation of diversification shifts within this group that, of course, may have occurred more recently.

Discussion

Our analyses provide strong evidence that although the cactus lineage is of moderate age, most of the extant diversity in this group was generated by significant radiations occurring throughout the mid to late Miocene and into the Pliocene. The timing of these major diversification events within cacti is extraordinarily similar to those inferred for other more distantly related succulent plant groups. Agaves, with their center of diversity in North American deserts, are reported to have had two primary pulses of rapid diversification, the first at 8–6 Ma and a second at 3–2.5 Ma (Good-Avila, et al. 2006). In South Africa, the ice plants (particularly the core Ruschioideae) comprise the fundamental element of the succulent Karoo region, and their major radiation was estimated as occurring roughly 8.7–3.8 Ma (Klak, et al. 2004). Remarkably, new research in *Euphorbia* has identified many multiple independent radiations of succulent lineages, occurring across regions of Africa,

Madagascar, and South America, and all within a timeframe of ≈11–2 Ma (Horn 2010). While we currently lack data on the aloes, stonecrops, and asclepiads, every arid-adapted succulent lineage investigated thus far has followed a similar tempo of evolution: Although the origins of a pronounced succulent syndrome in these groups vary widely between ≈40 and 10 Ma, they all share a single timeframe of extensive global diversification in the late Miocene-Pliocene. Studies of other desert (nonsucculent) plants from various regions have also demonstrated a similar pattern of recent radiation (Moore and Jansen 2006, Catalano, et al. 2008, Luebert and Wen 2008, Verboom, et al. 2009). The simultaneous diversification of arid-adapted lineages provides a general insight into the history of the world's arid regions, which has been limited by a bias against fossil formation in dry environments. As a unique paleoclimate proxy, this timeframe is in good agreement with other evidence that the late Miocene-Pliocene witnessed the establishment of many extant desert ecosystems.

Most of the succulent radiations reported here could be reasonably linked to the expansion of aridity due to particular geological events. In North American cacti and agaves for example, radiations coincided with the establishment of the Sonoran desert, which was presumably caused by increased volcanic activity and the formation of the Gulf of California (Ferrusquía-Villafranca and Gonzáles-Guzmán 2005). This timeframe was also a period of significant Andean uplift activity, which both intensified and expanded arid environments throughout most of western South America (Gregory-Wodzicki 2000). Eastern Africa was similarly experiencing increasing aridity possibly due to shifts in ocean (Cane and Molnar 2001) or atmospheric (Sepulchre, et al. 2006) circulation, and around this time a winter rainfall precipitation regime became established in the South African Karoo region (Klak, et al. 2004, Verboom, et al. 2009).

However, the temporal concordance of these diversification events with major vegetation changes in other geographical regions suggests that a more global environmental driver may also be contributing to the expansion of drought-adapted plant communities (Fig. 2). The late Miocene has long been recognized as a fundamental moment in Earth history due to the global emergence of grasslands dominated by grasses with C_4 photosynthesis (Cerling, et al. 1997). C_4 photosynthesis is a highly convergent and complex trait that reduces rates of photorespiration. It is advantageous under low atmospheric CO_2 conditions and is frequently associated with plants adapted to environments that promote high levels of photorespiration, such as hot temperatures, aridity, and high salinity (Sage 2004). Although the earliest known origin of C_4 photosynthesis was ≈30 Ma and coincided with an abrupt CO_2 decline in the Oligocene, most origins appeared much later (Edwards, et al. 2010). In fact, a global burst of evolution of the C_4 photosynthetic pathway in angiosperms, inferred both in grasses (Edwards, et al. 2010) and in eudicots (Christin, et al. 2011a), occurred during the late Miocene-Pliocene (Fig. 2). This very small window of time thus witnessed extraordinary changes in ecosystem properties worldwide: The emergence of C_4 grasslands, which cover 20–30% of the terrestrial land surface; the origins of many C_4 plant lineages, which are signature elements of most stress-adapted floras; and a pulse of rapid diversification in every major succulent plant lineage that has thus far been examined. We highlight midlate

Miocene trends in two potential environmental variables that could have provided strong ecological advantages to both succulent and C_4 syndromes.

First, a steep and steady decline in global temperatures after the mid-Miocene climate optimum is well documented from oxygen isotope records of deep-sea foraminifera (Zachos, et al. 2001), although the causes of this decline are still debated (Ruddiman 2010). While the relationship between temperature and precipitation is complex, drops in global temperatures may reduce global precipitation due to a slowdown of the hydrological cycle, and evidence from both paleoecological reconstructions (reviewed in Edwards, et al. 2010) and stable isotopes (e.g, Dettman, et al. 2001, Huang, et al. 2007) from multiple continents consistently indicate a general trend toward increasing aridity during this time. Thus, irrespective of any one particular geological event, the late Miocene-Pliocene succulent and C_4 plant revolution corresponded with what appears to be a near-global phenomenon of reduced precipitation.

Second, we argue that the coincidental nature of this global succulent diversification with the rise of C_4 photosynthesis warrants another look at atmospheric CO_2 levels during the late Miocene. There has been much disagreement about CO_2 concentrations during this time; although the abrupt Oligocene decline in atmospheric CO_2 is not debated, different CO_2 proxies have produced conflicting signals regarding subsequent fluctuation (Christin, et al. 2011a). Most recently, Tripati et al. (2009) estimated a precipitous decline in CO_2 between 15 and 8 Ma, from roughly 425 ppm to 200 ppm (Fig. 2), with further fluctuations between 200 and 300 ppm occurring into the Pleistocene. A drop of that magnitude would carry disastrous consequences for C_3 plants (Sage 2004) and would have provided a strong and obvious competitive advantage to the C_4 syndrome. These same trends could also promote diversification of lineages that already possess a suite of drought-adapted traits. Declining CO_2 decreases a typical plant's water use efficiency, because the diffusion gradient between atmospheric and internal leaf CO_2 levels will be smaller and plants will need to adopt a higher stomatal conductance (and thus greater water loss) to maintain a given rate of carbon fixation. A drop in CO_2 concentration would therefore immediately expand the ecological space in which drought-adapted succulent plants, with their high photosynthetic water use efficiency, would be competitive (Ehleringer and Monson 1993).

We suggest that a rapid expansion of available habitat (rather than any particular new “key innovation”) during the late Miocene was a primary driver of the global diversification of plant lineages already possessing a preadapted succulent syndrome. Against a backdrop of increasing global aridity, a sharp CO_2 decline is a plausible driver of the simultaneous expansion of C_4 grasslands, the clustering of new C_4 origins, and the diversification of succulent lineages. The contemporaneous spread of multiple C_4 and succulent plant lineages across the global landscape is a remarkable demonstration of convergence in plants, and their limited and predictable evolutionary responses to environmental stress.

Acknowledgments

We thank S. Albesiano, M. Alfaro, C. Blazier, J. Brown, F. Cáceres, A. Cano, D. Chatelet, C. Dunn, L. Garrison, F. Goetz, C. Hunkeler, D. Jarzen, R. Kiesling, S. Manchester, I. Peralta, A. Rahlin, D. Royer, S. Schmerler, S. Smith, D. Soltis, P. Soltis, V. Vincenzetti, and Z. Xi for assistance during plant collection and laboratory work, advice with analyses, and comments on the manuscript and the staff of various Herbarium/Botanical Gardens: BRU, FR, G, MERL, MO, NY, UPS, USM, Z, and ZSS (Table S2). This work was supported by National Science Foundation-Division of Environmental Biology Grant 1026611 (to E.J.E.).

References

Alfaro ME, Santini F, Brock C, Alamillo H, Dornburg A, Rabosky DL, Carnevale G, Harmon LJ. 2009. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. Proceedings of the National Academy of Sciences, 106:13410-13414.

Anderson EF. 2001. The cactus family. Portland, Oregon (U.S.A.), Timber press, Inc.

Axelrod D. 1979. Age and origin of sonoran desert vegetation. Occ Pap Cal Acad Sci, 132:1–74.

Baldwin BG, Sanderson MJ. 1998. Age and rate of diversification of the hawaiian silversword alliance (compositae). Proc. Natl. Acad. Sci. U. S. A., 95:9402-9406.

Barreda VD, Palazzesi L, Tellería MC, Katinas L, Crisci JV, Bremer K, Passalia MG, Corsolini R, Brizuela RR, Bechis F. 2010. Eocene patagonia fossils of the daisy family. Science, 329:1621.

Call VB, Dilcher DL. 1992. Investigations of angiosperms from the eocene of southeastern north-america - samaras of fraxinus-wilcoxiana berry. Rev. Palaeobot. Palynology, 74:249-266.

Cane MA, Molnar P. 2001. Closing of the indonesian seaway as a precursor to east african aridification around 3-4[thinsp]million years ago. Nature, 411:157-162.

Catalano SA, Vilardi JC, Tosto D, Saidman BO. 2008. Molecular phylogeny and diversification history of prosopis (fabaceae : Mimosoideae). Biol. J. Linnean Soc., 93:621-640.

Cerling TE, Harris JM, MacFadden BJ, Leakey MG, Quade J, Eisenmann V, Ehleringer jR. 1997. Global vegetation change through the miocene/pliocene boundary. Nature, 389:153-158.

Cevallos-Ferriz SRS, Estrada-Ruiz E, Perez-Hernandez BR. 2008. Phytolaccaceae infructescence from cerro del pueblo formation, upper cretaceous (late campanian), coahuila, mexico. American Journal of Botany, 95:77-83.

Chevreux B. 1997–2010. Mira v1-v3., Dept. of Molecular Biophysics, Deutsches Krebsforschungszentrum Heidelberg, Germany. .

Christin P-A, Osborne CP, Sage RF, Arakaki M, Edwards EJ. 2011a. C4 eudicots are not younger than c4 monocots. Journal of Experimental Botany, 62:3171-3181.

Christin PA, Besnard G, Samaritani E, Duvall MR, Hodkinson TR, Savolainen V, Salamin N. 2008. Oligocene co2 decline promoted c-4 photosynthesis in grasses. Curr. Biol., 18:37-43.

Christin PA, Sage TL, Edwards EJ, Ogburn RM, Khoshravesh R, Sage RF. 2011b. Complex evolutionary transitions and the significance of c-3-c-4 intermediate forms of photosynthesis in molluginaceae. Evolution, 65:643-660.

Collinson M, Boulter M, Holmes P. 1993. The fossil record 2. Chapman & Hall, London.

Crane PR, Pedersen KR, Friis EM, Drinnan AN. 1993. Early cretaceous (early to middle albian) platanoid inflorescences associated with sapindopsis leaves from the potomac group of eastern north america. Syst. Bot., 18:328-344.

Crepet W, Nixon K. 1998. Fossil clusiaceae from the late cretaceous (turonian) of new jersey and implications regarding the history of bee pollination. American Journal of Botany, 85:1122.

Crepet WL, Feldman GD. 1991. The earliest remains of grasses in the fossil record. American Journal of Botany, 78:1010-1014.

Crepet WL, Nixon KC, Gandolfo MA. 2004. Fossil evidence and phylogeny: The age of major angiosperm clades based on mesofossil and macrofossil evidence from cretaceous deposits. American Journal of Botany, 91:1666-1682.

Dettman DL, Kohn MJ, Quade J, Ryerson FJ, Ojha TP, Hamidullah S. 2001. Seasonal stable isotope evidence for a strong asian monsoon throughout the past 10.7 m.Y. Geology, 29:31-34.

Edgar R. 2004. Muscle: A multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics, 5:113.

Edwards EJ, Donoghue MJ. 2006. Pereskia and the origin of the cactus life-form. Am. Nat., 167:777-793.

Edwards EJ, Nyffeler R, Donoghue MJ. 2005. Basal cactus phylogeny: Implications of pereskia (cactaceae) paraphyly for the transition to the cactus life form. American Journal of Botany, 92:1177-1188.

Edwards EJ, Osborne CP, Strömberg CAE, Smith SA, Consortium CG. 2010. The origins of c4 grasslands: Integrating evolutionary and ecosystem science. Science, 328:587-591.

Ehleringer JR, Monson RK. 1993. Evolutionary and ecological aspects of photosynthetic pathway variation. Annu. Rev. Ecol. Syst., 24:411-439.

Ferrusquía-Villafranca I, Gonzáles-Guzmán L. 2005. Biodiversity, ecosystems, and conservation in northern mexico. Oxford Univ Press, New York.

Friis E, Crane P, Pedersen K. 1998. The reproductive structures of cretaceous platanaceae. Biologiske Skrifter, 31:1–55.

Friis EM, Eklund H, Pedersen KR, Crane PR. 1994. Virginianthus calycanthoides gen et sp-nov - a calycanthaceous flower from the potomac group (early cretaceous) of eastern north-america. Int. J. Plant Sci., 155:772-785.

Gibson AC, Nobel PS. 1986. The cactus primer. Cambridge (US:MA)/London (GB), Harvard University Press.

Good-Avila SV, Souza V, Gaut BS, Eguiarte LE. 2006. Timing and rate of speciation in agave (agavaceae). Proceedings of the National Academy of Sciences, 103:9124-9129.

Gregory-Wodzicki KM. 2000. Uplift history of the central and northern andes: A review. Geol. Soc. Am. Bull., 112:1091-1105.

Hershkovitz MA, Zimmer EA. 1997. On the evolutionary origins of the cacti. Taxon, 46:217-231.

Horn J, et al. 2010. Are growth forms and photosynthetic pathways correlates of diversification in euphorbia (euphorbiaceae)? Botany 2010, p. 91.

Huang Y, Clemens SC, Liu W, Wang Y, Prell WL. 2007. Large-scale hydrological change drove the late miocene c4 plant expansion in the himalayan foreland and arabian peninsula. Geology, 35:531-534.

Hughes N. 1994. The enigma of angiosperm origins. Cambridge, UK, Cambridge Univ Press.

Hunt DR, Taylor N, Graham C. 2006. The new cactus lexicon. Dorchester (UK), dh books.

Jansen RK, Raubeson LA, Boore JL, DePamphilis CW, Chumley TW, Haberle RC, Wyman SK, Alverson AJ, Peery R, Herman SJ*, et al.* 2005. Methods for obtaining and analyzing whole chloroplast genome sequences. In: Zimmer EA, Roalson EH editors. Molecular evolution: Producing the biochemical data, part b. San Diego, Elsevier Academic Press Inc, p. 348-384.

Klak C, Reeves G, Hedderson T. 2004. Unmatched tempo of evolution in southern african semi-desert ice plants. Nature, 427:63-65.

Knobloch E, Mai D. 1986. Monographie der früchte und samen in der kreide von mitteleuropa. Rozpravy ústredního ústavu geologickénho Praha, 47:1-219.

Linder HP. 1987. The evolutionary history of the poales/restionales: A hypothesis. Kew Bulletin, 42:297-318.

Luebert F, Wen J. 2008. Phylogenetic analysis and evolutionary diversification of heliotropium sect. Cochranea (heliotropiaceae) in the atacama desert. Syst. Bot., 33:390-402.

Manchester S, O'Leary E. 2010. Phylogenetic distribution and identification of fin-winged fruits. The Botanical Review, 76:1-82.

Mapes G, Rothwell GW. 1984. Permineralized ovulate cones of lebachia from late paleozoic limestones of kansas. Paleontology, 27:69-94.

Mathews S, Lavin M, Sharrock R. 2005. Evolution of the phytochrome gene family and its utility for phylogenetic analyses of angiosperms. Ann Mo Bot Gard, 82:296–321.

Moore MJ, Bell CD, Soltis PS, Soltis DE. 2007. Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. Proc. Natl. Acad. Sci. U. S. A., 104:19363-19368.

Moore MJ, Jansen RK. 2006. Molecular evidence for the age, origin, and evolutionary history of the american desert plant genus tiquilia (boraginaceae). Mol. Phylogenet. Evol., 39:668-687.

Moore MJ, Soltis PS, Bell CD, Burleigh JG, Soltis DE. 2010. Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. Proc. Natl. Acad. Sci. U. S. A., 107:4623-4628.

Nichols DJ, Traverse A. 1971. Palynology, petrology, and depositional environments of some early tertiary lignites in texas. Geoscience and Man, 3:37-48.

Nyffeler R. 2002. Phylogenetic relationships in the cactus family (cactaceae) based on evidence from trnk/matk and trnl-trnf sequences. American Journal of Botany, 89:312-326.

Nyffeler R, Eggli U. 2010a. Disintegrating portulacaceae: A new familial classification of the suborder portulacineae (caryophyllales) based on molecular and morphological data. Taxon, 59:227-240.

Nyffeler R, Eggli U. 2010b. A farewell to dated ideas and concepts: Molecular phylogenetics and a revised suprageneric classification of the family cactaceae. Schumannia, 6:109-149.

Nyffeler R, Eggli U, Ogburn M, Edwards E. 2008. Variations on a theme: Repeated evolution of succulent life forms in the portulacineae (caryophyllales). Haseltonia:26-36.

O'Quinn R, Hufford L. 2005. Molecular systematics of montieae (portulacaceae): Implications for taxonomy, biogeography and ecology. Syst. Bot., 30:314-331.

Ocampo G, Columbus JT. 2010. Molecular phylogenetics of suborder cactineae (caryophyllales), including insights into photosynthetic diversification and historical biogeography. American Journal of Botany, 97:1827-1847.

Ogburn RM, Edwards EJ. 2010. The ecological water-use strategies of succulent plants. In: Kader JC, Delseny M editors. Advances in botanical research, vol 55. London, Academic Press Ltd-Elsevier Science Ltd, p. 179-225.

Pedersen KR, Friis EM, Crane PR, Drinnan AN. 1994. Reproductive structures of an extinct platanoid from the early cretaceous (latest albian) of eastern north-america. Rev. Palaeobot. Palynology, 80:291-303.

Prasad V, Strömberg CAE, Alimohammadian H, Sahni A. 2005. Dinosaur coprolites and the early evolution of grasses and grazers. Science, 310:1177-1180.

Ruddiman WF. 2010. A paleoclimatic enigma? Science, 328:838-839.

Rutschmann F. 2006. Molecular dating of phylogenetic trees: A brief review of current methods that estimate divergence times. Divers. Distrib., 12:35-48.

Sage RF. 2004. The evolution of c-4 photosynthesis. New Phytol., 161:341-370.

Sepulchre P, Ramstein G, Fluteau F, Schuster M, Tiercelin J-J, Brunet M. 2006. Tectonic uplift and eastern africa aridification. Science, 313:1419-1423.

Sims HJ, Herendeen PS, Lupia R, Christopher RA, Crane PR. 1999. Fossil flowers with normapolles pollen from the upper cretaceous of southeastern north america. Rev. Palaeobot. Palynology, 106:131-151.

Stamatakis A. 2006. Raxml-vi-hpc: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics, 22:2688-2690.

Strömberg CAE. 2005. Decoupled taxonomic radiation and ecological expansion of open-habitat grasses in the cenozoic of north america. Proc. Natl. Acad. Sci. U. S. A., 102:11980-11984.

Thorne JL, Kishino H. 2002. Divergence time and evolutionary rate estimation with multilocus data. Syst. Biol., 51:689-702.

Thorne JL, Kishino H, Painter IS. 1998. Estimating the rate of evolution of the rate of molecular evolution. Mol. Biol. Evol., 15:1647-1657.

Tripati AK, Roberts CD, Eagle RA. 2009. Coupling of co2 and ice sheet stability over major climate transitions of the last 20 million years. Science, 326:1394-1397.

Verboom GA, Archibald JK, Bakker FT, Bellstedt DU, Conrad F, Dreyer LL, Forest F, Galley C, Goldblatt P, Henning JF*, et al.* 2009. Origin and diversification of the greater cape flora: Ancient species repository, hot-bed of recent radiation, or both? Mol. Phylogenet. Evol., 51:44-53.

Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with dogma. Bioinformatics, 20:3252-3255.

Yang Z. 1993. Maximum-likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. Mol. Biol. Evol., 10:1396-1401.

Zachos J, Pagani M, Sloan L, Thomas E, Billups K. 2001. Trends, rhythms, and aberrations in global climate 65 ma to present. Science, 292:686-693.

Figures and Tables

Figure1. Time-calibrated phylogeny of the cacti and their relatives. Colored branches indicate shifts in diversification: Blue branches represent lineages with significantly lower net diversification than the background rate; green, orange, and pink branches indicate higher diversification and/or species turnover (see Table 1 for parameter estimates and clade names). Gray boxes indicate ecologically important succulent clades: Cactaceae (New World); Malagasy Didiereaceae (Madagascar); core Ruschiodeae (Aizoaceae, Southern Africa).

Figure 2. CO_2, global temperature, C_4 origins, C_4 grasslands, and the diversification of succulent plants during the late Miocene/Pliocene. Lines extend back to the origin of the various succulent clades, and significant diversification events are represented by increases in line width. For the ice plants, dark green indicates timing of diversification by Klak et al. (2004), and light green represents our estimated age of the same node ("core Rushioideae"; (Klak, et al. 2004). Blue line reflects decline in relative global temperatures, inferred from deep sea ^{18}O, which is primarily a metric of deep sea temperature and sea-ice volume. Gray area in background represents reconstructed atmospheric CO_2 levels and their uncertainty through time, collated from multiple proxies (Edwards, et al. 2010). Black line is the drop in CO_2 hypothesized by Tripati et al. (Tripati, et al. 2009).

Table 1. Significant shifts in diversification rate and species turnover in cacti and relatives

Supporting information

Figure S1. Chronogram of Angiosperms. Divergence times were estimated with Multidivtime on a ML tree of 102 taxa and 83 chloroplast genes. Clade in orange includes additions to analysis by Moore et al. (Moore, et al. 2010) to increase sampling in the Caryophyllales. Dots indicate placement of fossil constraints applied to dating analyses. Bootstrap values are above the branches. The root is not shown.

Figure S2. Maximum likelihood phylogram based on sequences of nuclear PHYC and chloroplast trnK/matK genes for 295 taxa representing the Portulacineae and outgroups. Black dots indicate placement of age constraints in secondary dating analyses. Numbers above the branches are likelihood bootstrap support values.

Figure S3. Calibrated tree from Fig. 1, pruned to representative taxa for MEDUSA analyses. The number of species that each taxon in the tree represents is placed alongside tip names. Outgroups were excluded from MEDUSA analyses because of lower representation of taxa.

Figure 1. Time-calibrated phylogeny of the cacti and their relatives

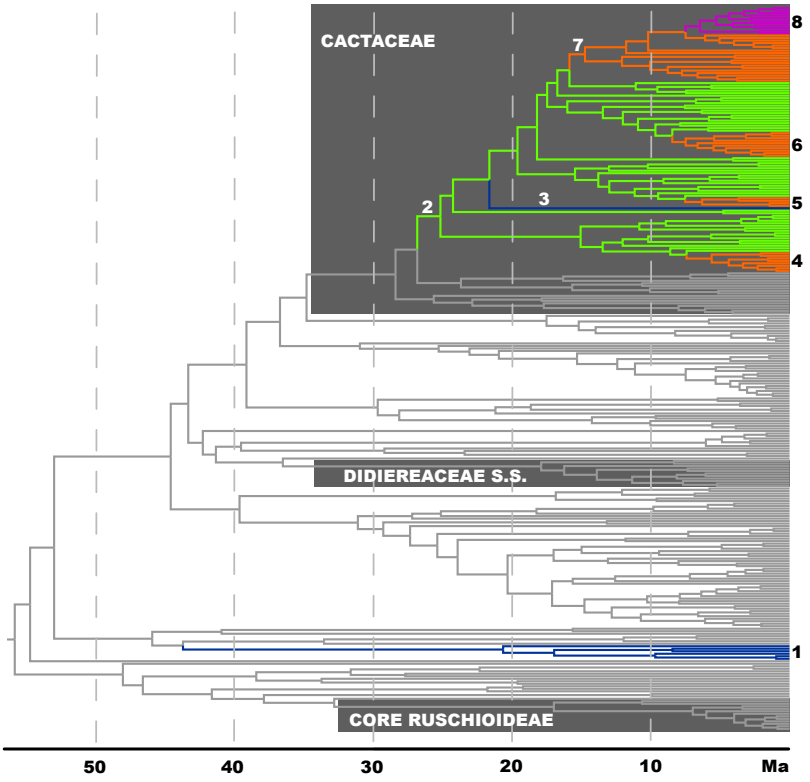


Figure 2. CO2, global temperature, C4 origins, C4 grasslands, and the diversification of succulent plants during the late Miocene/Pliocene

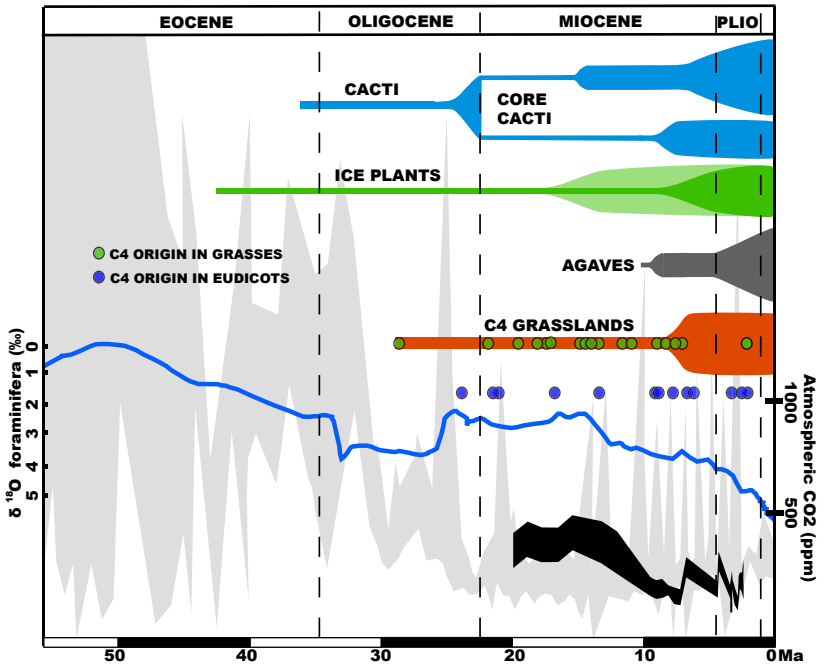


Table 1. Significant shifts in diversification rate and species turnover in cacti and relatives

Node number (see Fig. 1)	Clade	Age, Ma	*D*, lineage/Ma	*r*, lineage/Ma	ε	Center of diversity
Background rate	Portulacineae + Molluginaceae	55–53	0.137	0.095	1.0×10^{-7}	Worldwide
1	Molluginaceae pro parte	44–21	0.088	0.016	0.98	Southern Africa
2	Core cacti	27–25	0.268	0.232	0.306	Widespread in North and South America
3	*Blossfeldia liliputana*	21–0	0	2.27×10^{-17}	3.90×10^{-6}	South America
4	*Opuntia*	7.5–0	0.7	0.434	0.874	North America
5	*Mammillaria* + *Coryphantha*	7.6–6.3	0.719	0.225	0.973	Mexico
6	Hylocereinae + Echinocereinae	8.5–7.5	0.576	0.422	0.724	Mexico and Central America
7	Notocacteae + Cereeae	16.0–14.8	0.386	0.239	0.85	South America
8	Trichocereinae + Cereinae	7.5–6.5	0.768	0.432	0.903	South America

D is diversification rate under a pure-birth model (D = [ln(Nt) – ln(No)]/T, where T is the stem age of the clade, Nt is the number of taxa, and No = 1) (71). r is diversification rate (λ − μ) and ε is a calculation of species turnover rate (μ/λ) where λ = speciation rate and μ = extinction rate, as estimated by MEDUSA (9).

Figure S1. Chronogram of Angiosperms

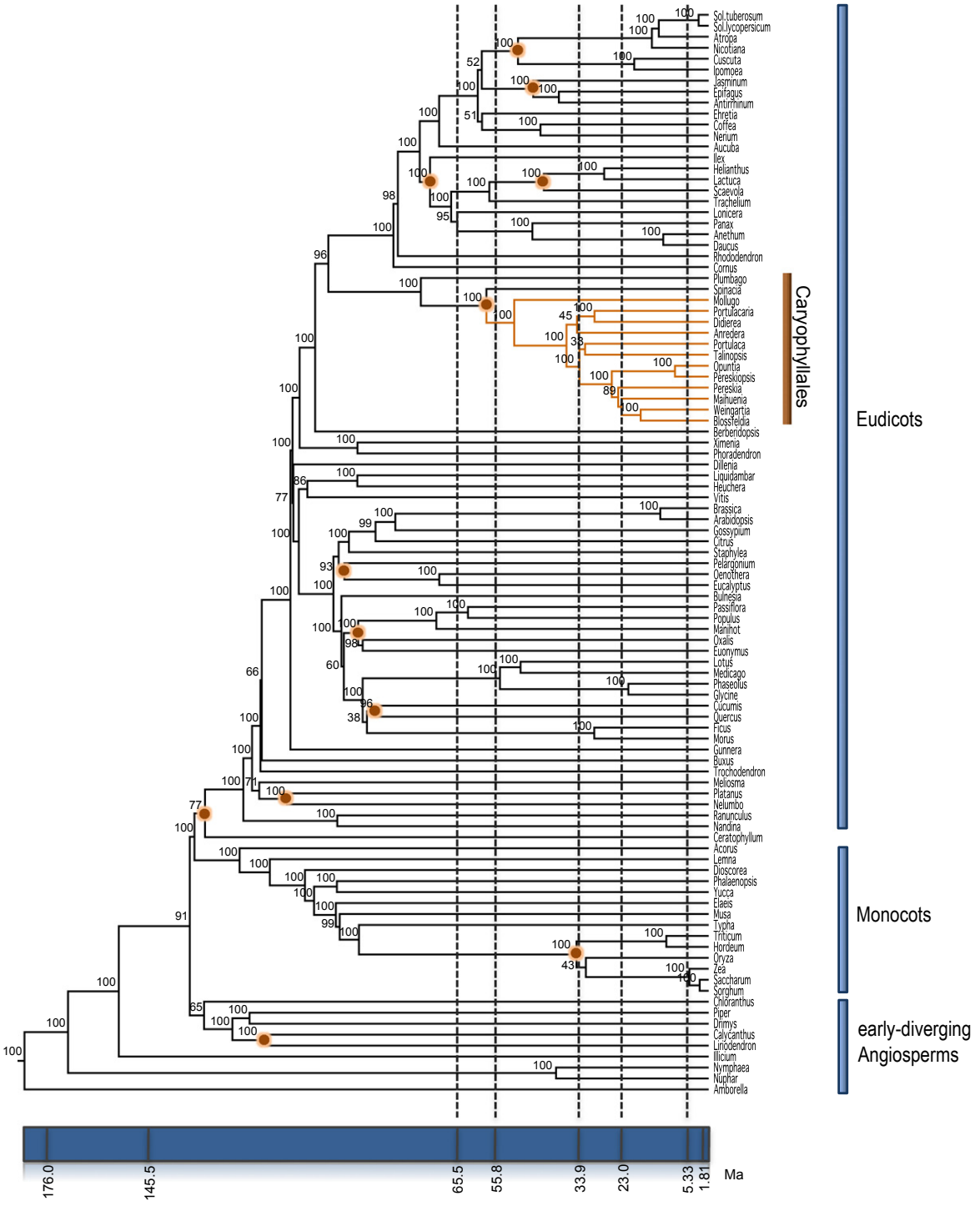


Figure S2. Maximum likelihood phylogram

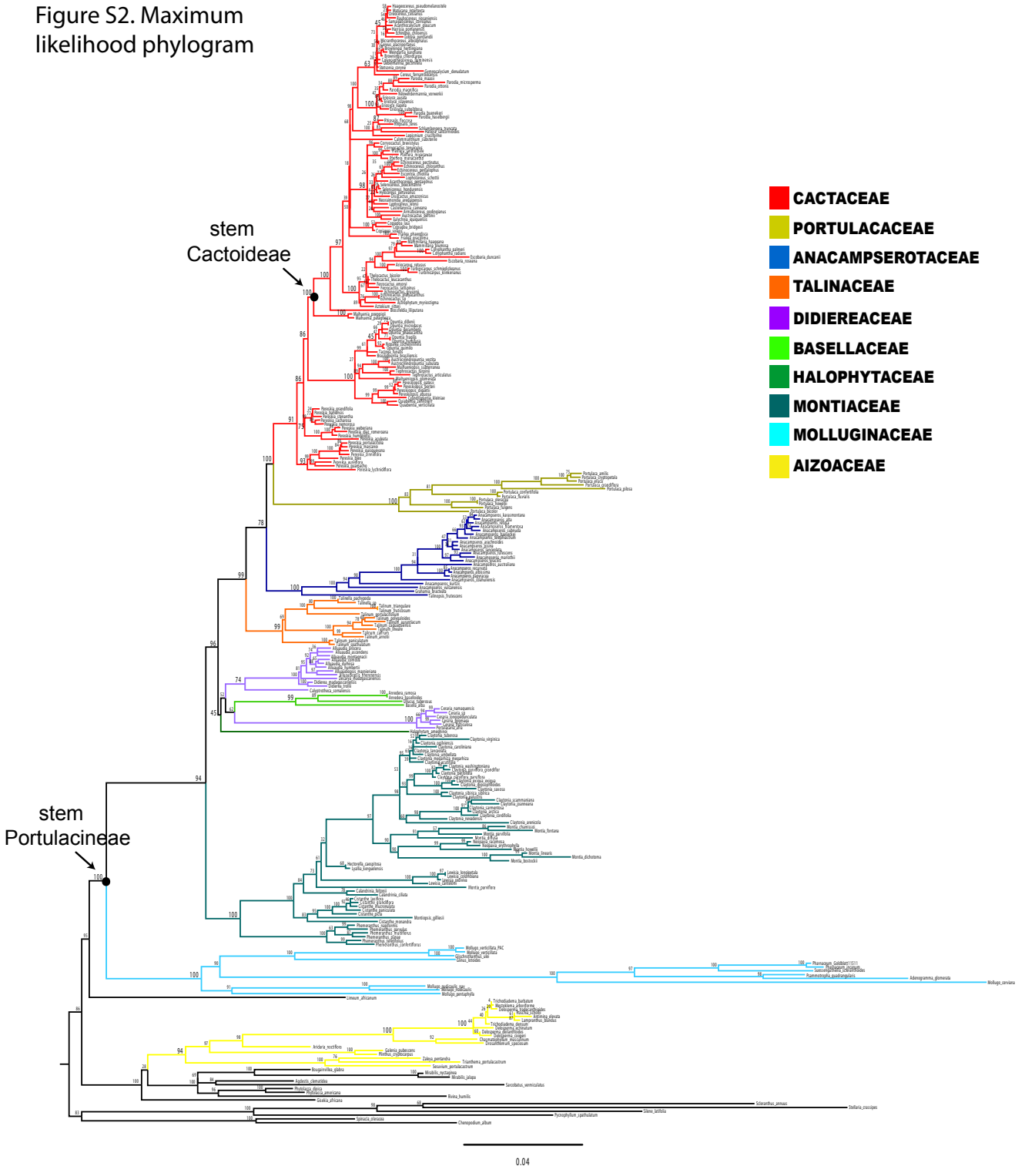


Figure S3 Calibrated phylogeny of the cacti and their relatives (MEDUSA representatives)

Table S1. Voucher information, and GenBank accession numbers* for species in the Caryophyllales added to 83-gene chloroplast dataset of Moore *et al.* (8).

Taxon	Provenance
Anredera baselloides Kunth	Ogburn 256, cultivated at Brown Univ.
Blossfeldia liliputana Werderm.	Nyffeler n.a., cultivated
Didierea madagascariensis Baill.	Rauh M1803, cultivated at ZSS (931246/0)
Maihuenia poeppigii (Otto & Pfeiff.) F.A.C. Weber ex K. Schum.	Hahn n.a., cultivated at ZSS (993081/0)
Mollugo verticillata L.	Arakaki 1892a, Brown Univ. campus.
Opuntia decumbens Salm-Dyck	Martínez & Eggli 146a, cultivated at ZSS (932161b)
Pereskia saccharosa Griseb.	Ritter 640, cultivated at ZSS (922123/0)
Pereskiopsis diguetii (F.A.C. Weber) Britton & Rose	Lomelí & Díaz n.a., cultivated at ZZS (942160/b)
Portulaca oleracea L.	Nyffeler n.a., Botanical Garden Zürich
Portulacaria afra Jacq.	Without collector inf., cultivated at ZSS (995326/0)
Talinopsis frutescens A. Gray	Ferguson 1352, cultivated at ZSS (931203/0)
Weingartia kargliana Rausch	Rausch 677, cultivated at ZSS (782218/b)

* *accD*: HQ620718-HQ620728, *atpA*: HQ620729-HQ620740, *atpB*: HQ620741-HQ620750, *atpE*: HQ620751-HQ620758, *atpF*: HQ620759-HQ620769, *atpH*: HQ620770-HQ620780, *atpI*: HQ620781-HQ620790, *ccsA*: HQ620791-HQ620801, *cemA*: HQ620802-HQ620812, *clpP*: HQ620813-HQ620818, *infA*: HQ620819-HQ620829, *matK*: HQ620830-HQ620837, *ndhA*: HQ620897-HQ620905, *ndhB*: HQ620906-HQ620914, *ndhC*: HQ620915-HQ620921, *ndhD*: HQ620922-HQ620932, *ndhE*: HQ620933-HQ620941, *ndhF*: HQ620942-HQ620949, *ndhG*: HQ620950-HQ620958, *ndhH*: HQ620959-HQ620968, *ndhI*: HQ620969-HQ620977, *ndhJ*: HQ620978-HQ620984, *ndhK*: HQ620985-HQ620992, *petA*: HQ620993-HQ621002, *petB*: HQ621003-HQ621013, *petD*: HQ621014-HQ621024, *petG*: HQ621025-HQ621035, *petL*: HQ621036-HQ621044, *petN*: HQ621045-HQ621054, *psaA*: HQ621149-HQ621158, *psaB*: HQ621159-HQ621168, *psaC*: HQ621169-HQ621179, *psaI*: HQ621180-HQ621189, *psaJ*: HQ621190-HQ621197, *psbA*: HQ621198-HQ621208, *psbB*: HQ621209-HQ621218, *psbC*: HQ621219-HQ621228, *psbD*: HQ621686-HQ621695, *psbE*: HQ621229-HQ621239, *psbF*: HQ621240-HQ621249, *psbH*: HQ621250-HQ621260, *psbI*: HQ621261-HQ621269, *psbJ*: HQ621270-HQ621279, *psbK*: HQ621280-HQ621291, *psbL*: HQ621292-HQ621301, *psbM*: HQ621302-HQ621310, *psbN*: HQ621311-HQ621321, *psbT*: HQ621322-HQ621332, *rbcL*: HQ621333-HQ621341, *rpl2*: HQ621342-HQ621351, *rpl14*: HQ621352-HQ621362, *rpl16*: HQ621363-HQ621373, *rpl20*: HQ621374-HQ621384, *rpl22*: HQ621385-HQ621393, *rpl23*: HQ621394-HQ621402, *rpl32*: HQ621403-HQ621412, *rpl33*: HQ621413-HQ621423, *rpl36*: HQ621424-HQ621434, *rpoA*: HQ621435-HQ621445, *rpoB*: HQ621446-HQ621455, *rpoC1*: HQ621456-HQ621465, *rpoC2*: HQ621466-HQ621475, *rps2*: HQ621476-HQ621485, *rps3*: HQ621486-HQ621495, *rps4*: HQ621496-HQ621506, *rps7*: HQ621507-HQ621517, *rps8*: HQ621518-HQ621528, *rps11*: HQ621529-HQ621539, *rps12*: HQ621540-HQ621550, *rps14*: HQ621551-HQ621560, *rps15*: HQ621561-HQ621571, *rps16*: HQ621572-HQ621582, *rps18*: HQ621583-HQ621591, *rps19*: HQ621592-HQ621602, *rrn4.5*: HQ621603-HQ621613, *rrn5*: HQ621614-HQ621624, *rrn16*: HQ621625-HQ621635, *rrn23*: HQ621636-HQ621646, *ycf1*: HQ621647-HQ621654, *ycf2*: HQ621655-HQ621664, *ycf3*: HQ621665-HQ621674, *ycf4*: HQ621675-HQ621685.

Table S2. List of taxa, voucher information, and GenBank accession numbers for markers included in the present study.

Taxon	Voucher	GenBank acc. No. *trnk/matK*	*PHYC*
Aizoaceae – Aizooideae			
Galenia pubescens (Eckl. & Zeyh.) Druce		AY042589*	
Plinthus cryptocarpus Fenzl		AY042633*	
Aizoaceae – Mesembryanthemoideae			
Aridaria noctiflora (L.) N. E. Br.		AY042619*	
Aizoaceae – Ruschioideae			
Antimima elevata H.E.K. Hartmann	100171/0 (ZSS)	HQ620852	
Chasmatophyllum musculinum Dinter & Schwantes	994202/0 (ZSS)	HQ620857	
Delosperma cooperi L. Bolus		DQ855843*	
Delosperma deilanthoides S.A. Hammer	99 8369/0 (ZSS)	HQ620866	
Delosperma echinatum (Lam.) Schwantes		AY042575*	
Delosperma tradescantioides L. Bolus	995798/0 (ZSS)	HQ620867	
Drosanthemum speciosum Schwantes	931409/0 (ZSS)	HQ620868	
Lamprantus blandus Schwantes		FN597631*	
Mestoklema arboriforme (N.E.Burch)	824016/0 (ZSS)	HQ620879	
Ruschia schollii Schwantes		AY042649*	
Trichodiadema barbatum Schwantes		AY042666*	
Trichodiadema densum Schwantes	931431/0 (ZSS)	HQ620893	
Aizoaceae – Sesuvioideae			
Sesuvium portulacastrum (L.) L.	Sage 0509-07	FN868302*	
Trianthema portulacastrum L.	0010151 (FR)	FN825774*	
Zaleya pentandra (L.) C. Jeffrey	Sage 0509-04	FN868303*	
Amaranthaceae			
Chenopodium album L.		GQ434131*	
Spinacia oleracea L.		NC002202 *	
Anacampserotaceae			
Anacampseros alta Poelln. cf	901679/0 (ZSS)	HQ620844	HQ621064
Anacampseros albissima (Marloth) G.D. Rowley	Ogburn 255 (BRU)	DQ855856*	HQ621079
Anacampseros arachnoides Sims	892106/0 (ZSS)	HQ620845	HQ621065
Anacampseros australiana J.M. Black		DQ855855*	
Anacampseros baeseckei Dinter cf	893409/b (ZSS)	HQ620846	HQ621066
Anacampseros coahuilensis (S. Watson) Eggli & Nyffeler	901679/0 (ZSS)	AY875374*	HQ621067
Anacampseros filamentosa Sims	Ogburn 36 (MO)	HQ620847	HQ621068
Anacampseros gracilis Poelln.	901695/a (ZSS)	HQ620848	HQ621069
Anacampseros karasmontana Dinter ex Poelln.		DQ855859*	
Anacampseros kurtzii Bacigalupo	Leuenberger 4217 (B)	DQ855853*	HQ621070
Anacampseros lanceolata Sweet	Ogburn 38 (MO)		HQ621071
Anacampseros marlothii Poelln.	901684/b (ZSS)	HQ620849	HQ621072
Anacampseros papyracea (E. Mey. ex Fenzl) G.D. Rowley		DQ855857*	
Anacampseros pisina G. Will.	997039/0 (ZSS)	HQ620850	
Anacampseros retusa Poelln.	803518/a (ZSS)	DQ855860*	HQ621073

Anacampseros recurvata (Schönland) G.D. Rowley		DQ855858*	
Anacampseros rufescens Sweet	Ogburn 37 (MO)		HQ621074
Anacampseros subnuda Poelln.		DQ855861*	
Anacampseros telephiastrum DC.	901682/a (ZSS)	AY875373*	HQ621075
Anacampseros vulcanensis Añon	904035/0 (ZSS)	AF542597*	HQ621076
Grahamia bracteata Gillies ex Hook.	B 142-32-94-10	AY015273*	AY875308*
Talinopsis frutescens A. Gray	Ogburn 31 (MO)	DQ855851*	HQ621135
Basellaceae			
Anredera baselloides (Kunth) Baill.	Ogburn 256 (BRU)	HQ620830	HQ621077
Anredera ramosa (Moq.) Eliasson	981287/0 (ZSS)	HQ620851	
Basella alba L.		AY042553*	
Ullucus tuberosus Caldas	101123/0 (ZSS)	HQ620896	HQ621147
Cactaceae – Cactoideae			
Acanthocalycium glaucum F. Ritter		AY015325*	
Acanthocereus pentagonus (L.) Britton & Rose		AY015295*	
Ariocarpus retusus Scheidw.	Arakaki 1896 (BRU)	HQ620853	
Armatocereus godingianus (Britton & Rose) Backeb. ex E. Salisb.	AY015296*		
Astrophytum myriostigma Lem.		AY015288*	
Austrocactus bertinii (E. Cels) Britton & Rose		AY015300*	
Aztekium ritteri (Boed.) Boed.	862607 (ZSS)	AY015290*	HQ621080
Blossfeldia liliputana Werderm.	Nyffeler n.a., cultivated		AY875301*
Browningia chlorocarpa (Kunth) W.T. Marshall		AY015316*	
Browningia hertlingiana (Rauh) Buxb.		AY015315*	
Calymmanthium substerile F. Ritter		AY015291*	AY875314*
Castellanosia caineana Cárdenas		AY015298*	
Cereus alacriportanus Pfeiff.		AY015313*	
Cereus fernambucensis Lem.	B 166-88-83-10		AY875293*
Coleocephalocereus fluminensis Backeb.	cultivated	AY015318*	
Copiapoa bridgesii Backeb.		AY015293*	
Copiapoa laui Diers & Esteves		AY015294*	
Copiapoa solaris (F. Ritter) F. Ritter		AY015292*	
Corryocactus brevistylus (K. Schum. ex Vaupel) Britton & Rose		AY015302*	
Corryocactus tenuiculus (Rauh & Backeb.) Hutchison		AY015303*	
Coryphantha palmeri Britton & Rose	Arakaki 1897 (BRU)	HQ620863	
Coryphantha radians (DC.) Britton & Rose	Arakaki 1898 (BRU)	HQ620864	
Disocactus amazonicus (K. Schum.) D.R. Hunt		AY015312*	
Echinocactus grusonii Hildm.	Ogburn 257 (BRU)	HQ620869	
Echinocactus platyacanthus Link & Otto	B 255-92-16-86	AY015287*	AY875294*
Echinocactus sp.	cultivated, no voucher inf.		HQ621097
Echinocereus chloranthus Rumpler	Arakaki 1899 (BRU)	HQ620870	
Echinocereus pectinatus (Scheidw.) Engelm.	Arakaki 1900 (BRU)	HQ620871	
Echinocereus pentalophus (DC.) Haage	912367 (ZSS)	AY015307*	
Echinopsis chiloensis (Colla) Britton & Rose		AY015322*	
Eriosyce aurata (Pfeiff.) Backeb.		AY015336*	
Eriosyce islayensis (Foerster) Katt.		AY015337*	
Eriosyce napina (Phil.) Katt.		AY015339*	
Eriosyce subgibbosa (Haw.) Katt.		AY015338*	
Escobaria duncanii (Hester) Backeb.	Arakaki 1901 (BRU)	HQ620872	
Escobaria roseana Buxb.	Arakaki 1902 (BRU)	HQ620873	
Escontria chiotilla (F.A.C. Weber) Rose		AY015308*	

Eulychnia iquiquensis (K. Schum.) Britton & Rose		AY015301*	
Ferocactus emoryi (Engelm.) Orcutt	Arakaki 1903 (BRU)	HQ620874	
Ferocactus latispinus (Haw.) Britton & Rose	Arakaki 1904 (BRU)	HQ620875	
Frailea gracillima (Lem.) Britton & Rose		AY015285*	
Frailea phaeodisca (Speg.) Backeb. & F.M. Knuth		AY015286*	
Gymnocalycium denudatum (Link & Otto) Pfeiff. ex Mittler		AY015317*	
Haageocereus pseudomelanostele (Werderm. & Backeb.) Backeb.		AY015329*	
Harrisia pomanensis (F.A.C. Weber ex K. Schum.) Britton & Rose		AY015324*	
Hatiora salicornioides (Haw.) Britton & Rose		AY015341*	
Hylocereus peruvianus Backeb.	911450 (ZSS)	AY015310*	
Lepismium cruciforme (Vell.) Miq.	882981 (ZSS)	AY015344*	
Leptocereus leonii Britton & Rose		AY015297*	
Lobivia pentlandii L. Bolus		AY015323*	
Lophocereus schottii (Engelm.) Britton & Rose	921332 (ZSS)	AY015309*	
Mammillaria haageana Pfeiff.		AY015289*	
Mammillaria plumosa F.A.C. Weber	Arakaki 1907 (BRU)	HQ620878	
Matucana intertexta F. Ritter		AY015327*	
Micranthocereus albicephalus (Buining & Brederoo) F. Ritter		AY015314*	
Neoraimondia arequipensis Backeb.		AY015299*	
Neowerdermannia vorwerkii Fric	99 8506 (ZSS)	AY015340*	
Oreocereus celsianus (Lem. ex Salm-Dyck) Riccob.	99 1180 (ZSS)	AY015328*	
Parodia buenekeri Buining		AY015331*	
Parodia haselbergii (F. Haage) F.H. Brandt		AY015330*	
Parodia maasii (Heese) A. Berger	211198 (ZSS)	AY015333*	HQ621111
Parodia magnifica (F. Ritter) F.H. Brandt		AY015332*	
Parodia microsperma (F.A.C. Weber) Speg.		AY015334*	
Parodia ottonis (Lehm.) N.P. Taylor	961515 (ZSS)	AY015335*	
Pfeiffera ianthothele (Monv.) F.A.C. Weber		AY015304*	
Pfeiffera miyagawae Barthlott & Rauh		AY015305*	
Pfeiffera monacantha (Griseb.) P.V. Heath		AY015306*	
Rauhocereus riosaniensis Backeb.		AY015326*	
Rhipsalis floccosa Salm-Dyck ex Pfeiff.	99 7050 (ZSS)	AY015342*	
Rhipsalis teres (Vell.) Steud.		AY042645*	
Samaipaticereus corroanus Cárdenas		AY015321*	
Schlumbergera truncata (Haw.) Moran		AY015343*	
Selenicereus boeckmannii (Otto ex Salm-Dyck) Britton & Rose		AY015311*	
Selenicereus hondurensis (K. Schum.) Britton & Rose	Arakaki 1910 (BRU)	HQ620889	HQ621133
Stetsonia coryne (Salm-Dyck) Britton & Rose		AY015320*	
Thelocactus bicolor (Galeotti) Britton & Rose	Arakaki 1911 (BRU)	HQ620891	
Thelocactus leucacanthus (Zucc. Ex Pfeiff.) Britton & Rose	Arakaki 1912 (BRU)	HQ620892	
Turbinicarpus klinkerianus Backeb. & W. Jacobsen	Arakaki 1913 (BRU)	HQ620894	
Turbinicarpus schmiedickeanus (Boed.) Buxb. & Backeb.	Arakaki 1914 (BRU)	HQ620895	
Uebelmannia pectinifera Buining	995036 (ZSS)	AY015319*	
Weingartia kargliana Rausch	782218/b (ZSS)	HQ620837	HQ621148
Cactaceae – Maihuenioideae			
Maihuenia patagonica (Phil.) Britton & Rose	B 013-30-81-10	AY015281*	AY875303*
Maihuenia poeppigii (Otto & Pfeiff.) F.A.C. Weber ex K. Schum.	B 048-15-93-10	AY015282*	AY875309*

Cactaceae – Opuntioideae			
Austrocylindropuntia subulata (Muehlenpf.) Backeb.	Ogburn 258 (BRU)	AY875364*	HQ621078
Austrocylindropuntia vestita (Salm-Dyck) Backeb.		AY015278*	
Brasiliopuntia brasiliensis (Willd.) A. Berger	B 153-76-7-480	AY875370*	HQ621081
Cylindropuntia kleiniae (DC.) F.M. Knuth aff	Ogburn 259 (BRU)	HQ620865	HQ621093
Maihueniopsis glomerata (Haw.) R. Kiesling	Arakaki 1905 (BRU)	HQ620877	HQ621101
Maihueniopsis subterranea (R.E. Fr.) E.F. Anderson		EU834746*	
Nopalea cochenillifera (L.) Salm-Dyck	Ogburn 260 (BRU)	HQ620881	HQ621109
Opuntia decumbens Salm-Dyck		HQ620833	
Opuntia dillenii (Ker Gawl.) Haw.	B 304-11-99-10	AY875369*	AY875302*
Opuntia fragilis (Nutt.) Haw.		EF590413*	
Opuntia humifusa (Raf.) Raf.		EF590414*	
Opuntia microdasys (Lehm.) Pfeiff.		AY042622*	
Opuntia phaeacantha Engelm.	Ogburn 261 (BRU)		HQ621110
Opuntia quimilo K. Schum.		AY015279*	
Pereskiopsis aquosa (F.A.C. Weber) Britton & Rose	Ogburn 262 (BRU)	HQ620882	HQ621112
Pereskiopsis diguetii (F.A.C. Weber) Britton & Rose	942160/b (ZSS)	AY015280*	HQ621113
Pereskiopsis gatesii Baxter	Ogburn 263 (BRU)	HQ620883	HQ621114
Pereskiopsis porteri (Brandegee ex F.A.C. Weber) Britton & Rose	B 169-03-84-30	HQ620884	HQ621115
Quiabentia verticillata (Vaupel) Vaupel ex Berger	B 154-12-74-80	AY042641*	HQ621131
Quiabentia zehntneri (Britton & Rose) Britton & Rose	B 163-09-88-30	AY875372*	HQ621132
Tacinga funalis Britton & Rose		AY042660*	
Tephrocactus articulatus (Pfeiff.) Backeb.	Ogburn 264 (BRU)	AY875367*	HQ621145
Tephrocactus turpinii (Lem.) Lem.	Edwards n.a. (BRU)		HQ621146
Cactaceae – Pereskioideae			
Pereskia aculeata Mill.		AY875355*	AY875312*
Pereskia aureiflora F. Ritter		AY875354*	AY875297*
Pereskia bahiensis Gürke		AY875351*	
Pereskia bleo (Kunth) DC.		AY875359*	AY875289*
Pereskia diaz-romeroana Cárdenas		AY875353*	
Pereskia grandifolia Haw.		AY875362*	AY875298*
Pereskia guamacho F.A.C. Weber		AY015275*	AY875291*
Pereskia humboldtii Britton & Rose aff		AY875356*	AY875287*
Pereskia lychnidiflora DC.		AY875358*	AY875286*
Pereskia marcanoi A.E. Areces-Mallea		AY875360*	AY875288*
Pereskia nemorosa Rojas Acosta		AY875350*	AY875296*
Pereskia portulacifolia (L.) DC.		AY875361*	AY875315*
Pereskia quisqueyana Alain		AY875352*	AY875292*
Pereskia sacharosa Griseb.		AY875363*	AY875299*
Pereskia stenantha F. Ritter		AY015276*	AY875295*
Pereskia weberiana K. Schum.		AY875357*	AY875313*
Pereskia zinniiflora DC.		AY015277*	AY875290*
Caryophyllaceae			
Pycnophyllum spathulatum Mattf.		FJ460220*	
Scleranthus annuus L.		FJ404869*	
Silene latifolia Poir.		EU749398*	
Stellaria crassipes Hultén		FJ404875*	
Didiereaceae			
Alluaudia ascendens (Drake) Drake	823558/0 (ZSS)	AY042541*	HQ621056

Alluaudia comosa (Drake) Drake	931255/0 (ZSS)	HQ620838	HQ621057
Alluaudia dumosa (Drake) Drake	931254/0 (ZSS)	HQ620839	HQ621058
Alluaudia humbertii Choux	996062/0 (ZSS)	HQ620840	HQ621059
Alluaudia montagnacii Rauh	811678/0 (ZSS)	HQ620841	HQ621060
Alluaudia procera (Drake) Drake	Ogburn 265 (BRU)	HQ620842	HQ621061
Alluaudiopsis fiherenensis Humbert & Choux	872404/0 (ZSS)	AY042542*	HQ621062
Alluaudiopsis marnieriana Rauh	812213/0 (ZSS)	HQ620843	HQ621063
Calyptrotheca somalensis Gilg		AY042563*	
Ceraria fruticulosa H. Pearson & Stephens	811470/0 (ZSS)	AY875371*	HQ621082
Ceraria longipedunculata Merxm. & Podlech	901396/0 (ZSS)	HQ620854	HQ621083
Ceraria namaquensis (Sond.) Pearson & Stephens	772320/0 (ZSS)	HQ620855	HQ621084
Ceraria pygmaea Pillans	843638/0 (ZSS)	HQ620856	HQ621085
Ceraria sp.	901392/0 (ZSS)		HQ621086
Decarya madagascariensis Choux	931246/0 (ZSS)	AY042574*	HQ621094
Didierea madagascariensis Baill.	931246/0 (ZSS)	HQ620831	HQ621095
Didierea trollii Capuron & Rauh	996013/0 (ZSS)	AY042576*	HQ621096
Portulacaria afra Jacq.	Ogburn 27 (MO)	AY875368*	HQ621129
Gisekiaceae			
Gisekia africana Kuntze		AY042591*	
Halophytaceae			
Halophytum ameghinoi (Speg.) Speg.	Edwards 255 (BRU)	AY042599*	HQ621099
Limeaceae			
Limeum africanum L.		AY042608*	
Molluginaceae			
Adenogramma glomerata Druce	Ogburn 142 (BRU)	FN825689*	HQ621055
Glinus lotoides L.	Errter 8859 (NY)	FN825692*	HQ621098
Glischrothamnus ulei Pilg.		FN825699*	
Mollugo cerviana (L.) Ser.	Reveal & Holmgrem 1968 (NY)	FN825710*	
Mollugo nudicaulis Lam.	Thulin & Bashir-Mohamed 6759 (UPS)	FN825730*	HQ621102
Mollugo nudicaulis var. *navassensis* Ekm.	Liogier 16585 (NY)	FN825726*	HQ621103
Mollugo pentaphylla L.	Nee 42741 (NY)	FN825734*	HQ621104
Mollugo verticillata L.	Edwards n.a. (BRU)	FJ404854*	HQ621105
Mollugo verticillata L.	Nee 37372 (G)	FN825743*	HQ621106
Pharnaceum Goldblatt11511		AY042629*	
Pharnaceum incanum L.	Ogburn 148 (BRU)	FN825748*	HQ621116
Psammotropha quadrangularis (L. f.) Fenzl	Ogburn 160 (BRU)	FN825755*	HQ621130
Suessenguthiella scleranthoides Friedr.		AY042659*	
Montiaceae			
Calandrinia ciliata (Ruiz & Pav.) DC.		AY764127*	
Calandrinia feltonii Skottsb.		AY042562*	
Cistanthe grandiflora (Lindl.) Carolin ex Hershkovitz	998914/0 (ZSS)	AY042568*	HQ621087
Cistanthe laxiflora (Phil.) Ford	Ogburn 167 (MERL)	HQ620858	HQ621088
Cistanthe monandra (Nutt.) Hershkovitz	Ogburn 137 (BRU)	HQ620859	HQ621089
Cistanthe mucronulata (Meyen) Ford	942055/0 (ZSS)	HQ620860	HQ621090
Cistanthe paniculata (Ruiz & Pav.) Carolin ex Hershkovitz	Arakaki 1894 (USM)	HQ620861	HQ621091
Cistanthe picta (Gill. ex Arn.) Carolin ex Hershkovitz	Ogburn 165 (MERL)	HQ620862	HQ621092

Claytonia acutifolia Pall. ex Willd.		AY764097*	
Claytonia arctica Adams		AY764096*	
Claytonia arenicola L.F. Hend.		AY764088*	
Claytonia caroliniana Michx.		AY764099*	
Claytonia cordifolia S. Watson		AY764100*	
Claytonia exigua Douglas ex Torr. & A. Gray		AY764089*	
Claytonia gypsophiloides Fisch. & C.A. Mey.		AY764090*	
Claytonia joanneana Roem. & Schult.		AY764101*	
Claytonia lanceolata Pursh		AY764102*	
Claytonia megarhiza (A. Gray) Parry ex S. Watson		AY764103*	
Claytonia nevadensis S. Watson		AY764104*	
Claytonia ogilviensis McNeill		AY764105*	
Claytonia palustris Swanson & Kelley		AY764106*	
Claytonia parviflora subsp. *grandiflora* John M. Miller & K.L. Chambers		AY764092*	
Claytonia parviflora Douglas ex Hook.		AY764093*	
Claytonia perfoliata Donn ex Willd.		AY764091*	
Claytonia sarmentosa C.A. Mey.		AY764107*	
Claytonia saxosa Brandegee		AY764094*	
Claytonia scammaniana Hultén		AY764108*	
Claytonia sibirica L.		AY764109*	
Claytonia tuberosa Pall. ex Willd.		AY764111*	
Claytonia umbellata S. Watson		AY764112*	
Claytonia virginica L.		AY764113*	
Claytonia washingtoniana (Suksd.) Suksd.		AY764095*	
Hectorella caespitosa Hook. f.		EF551350*	
Lewisia cantelovii J.T. Howell		AY042607*	
Lewisia columbiana (Howell ex A. Gray) B.L. Rob.		AY764126*	
Lewisia longipetala (Piper) S. Clay	cultivated, no voucher inf.	HQ620876	HQ621100
Lewisia rediviva Pursh		AY764125*	
Lyallia kerguelensis Hook. f.		EF551349*	
Montia bostockii (A.E. Porsild) S.L. Welsh		AY764114*	
Montia chamissoi (Ledeb. ex Spreng.) Greene		AY764120*	
Montia dichotoma (Nutt.) Howell		AY764115*	
Montia diffusa (Nutt.) Greene		AY764121*	
Montia fontana L.		AY764119*	
Montia howellii S. Watson		AY764117*	
Montia linearis (Douglas ex Hook.) Greene		AY764116*	
Montia parviflora Douglas	Ogburn n.a. (BRU)		HQ621107
Montia parvifolia (Moc. ex DC.) Greene		AY764122*	
Montiopsis gilliesii (Hook. & Arn.) D.I. Ford	Ogburn 166 (BRU)	HQ620880	HQ621108
Neopaxia erythrophylla P.B.Heenan		AY764123*	
Neopaxia racemosa (Buchanan) P.B.Heenan		AY764124*	
Phemeranthus confertiflorus (Greene) Hershkovitz	100095/0 (ZSS)	HQ620885	HQ621117
Phemeranthus multiflorus (Rose & Standl.) G. Ocampo		EU834747*	
Phemeranthus napiformis (DC.) Raf.	Ogburn 266 (BRU)		HQ621118
Phemeranthus parvulus (Rose & Standl.) T.M. Price	100096/0 (ZSS)		HQ621119
Phemeranthus punae (R.E. Fr.)	995811/a (ZSS)	EU834748*	HQ621120
Phemeranthus teretifolius (Pursh) Raf.		EU834749*	
Nyctaginaceae			
Bougainvillea glabra Choisy		AY042560*	
Mirabilis jalapa L.		AY042614*	
Mirabilis nyctaginea (Michx.) MacMill.	Ogburn n.a. (BRU)	AY042624*	

Phytolaccaceae			
Agdestis clematidea Moc. & Sessé ex DC.		AY042538*	
Phytolacca americana L.		DQ401362*	
Phytolacca dioica L.	20012142 (Z)	AY042631*	
Rivina humilis L.		AY514850*	
Portulacaceae			
Portulaca amilis Speg. aff	Ogburn 11 (MO)	HQ620886	HQ621121
Portulaca bicolor F. Muell.		DQ855848*	
Portulaca confertifolia Hauman	998912/0 (ZSS)	HQ620887	
Portulaca cryptopetala Speg.	998200/0 (ZSS)		HQ621122
Portulaca eruca Hauman	995251/0 (ZSS)	DQ855849*	HQ621123
Portulaca fluvialis D. Legrand	997376/0 (ZSS)	EU834750*	HQ621124
Portulaca fulgens Griseb.	995813/0 (ZSS)		HQ621125
Portulaca grandiflora Hook.	998914/0 (ZSS)	EU834751*	HQ621126
Portulaca howellii (D. Legrand) Eliasson	901308/b (ZSS)	HQ620888	
Portulaca oleracea L.	Ogburn 18 (MO)	AY875349*	HQ621127
Portulaca pilosa D. Legrand	Ogburn 10 (MO)		HQ621128
Sarcobataceae			
Sarcobatus vermiculatus (Hook.) Torr.		AY042652*	
Talinaceae			
Talinella pachypoda Eggli	Ogburn 22 (MO)	DQ855846*	HQ621134
Talinella sp.		AY514859*	
Talinum arnotii Hook. f.	932096/0 (ZSS)		HQ621136
Talinum aurantiacum Engelm.	Ogburn 267 (BRU)		HQ621137
Talinum caffrum Eckl. & Zeyh.	Ogburn 21 (MO)	AY042662*	HQ621138
Talinum caquaquiensis	961663/a (ZSS)		HQ621139
Talinum fruticosum (L.) Juss.		DQ855844*	
Talinum lineare Kunth		EU834752*	
Talinum paniculatum (Jacq.) Gaertn.	Ogburn 16 (MO)	AY015274*	HQ621140
Talinum polygaloides Gillies ex Arn.	941958/0 (ZSS)	DQ855845*	HQ621141
Talinum portulacifolium (Forssk.) Asch. ex Schweinf.	Ogburn 23 (MO)	DQ855847*	HQ621142
Talinum spathulatum Engelm. ex A. Gray	931207/0 (ZSS)	HQ620890	HQ621143
Talinum triangulare (Jacq.) Willd.	998911/0 (ZSS)	DQ855844*	HQ621144

*Sequences retrieved from GenBank.

Abbreviations: B (Berlin Botanical Garden), BRU (Brown University Herbarium), FR (Senckenberg Nature Museum), G (Geneva Botanical Gardens), MERL (Ruiz Leal Herbarium), MO (Missouri Botanical Garden), NY (New York Botanical Garden), UPS (Uppsala University), USM (San Marcos University Herbarium), Z (Zürich Botanical Garden), ZSS (Sukkulenten-Sammlung Zürich).

Table S3. Fossil calibrations applied to the dating analysis.

Fossil	Age (Ma)	Original Reference
Emporiaceae, *Emporia lockardii*, cones	290-310	Mapes & Rothwell (47)
Eudicots, unnamed pollen	125.0	Hughes (48)
Calycanthaceae, *Virginianthus*, flower	113-98	Friis *et al.* (49) Crepet *et al.* (50)
Platanaceae, several fossils:	113-98	
Aquia, *Platanocarpus*, inflorescence		Crane *et al.* (51)
Hamatia, inflorescence, flower		Pedersen *et al.* (52)
Platananthus, inflorescence		Friis *et al.* (53)
Combretaceae, *Dilcherocarpon combretoides*, fruit	112-93.5	Manchester & O'Leary (54)
Clusiaceae, *Paleoclusia*, flower and fruit	90.0	Crepet & Nixon (55)
Betulaceae, *Bedellia*, flower and fruit	88-83	Sims *et al.* (56)
Aquifoliaceae, *Ilex*, fruit	65.5	Knobloch & Mai (57)
Amaranthaceae, *Chenopodipollis multiplex*, pollen	65-56.5	Nichols & Traverse (58)
Solanaceae, *Cantisolanum daturoides*, fruit	55.8-33.9	Collinson *et al.* (59)
Oleaceae, *Fraxinus wilcoxiana*, fruit	55.8-33.9	Call & Dilcher (60)
Asteraceae, *Mutisiapollis*, inflorescence	48.6-40.4	Barreda *et al.* (61)
Poaceae, from phytolith analyses	34.0	Strömberg (62)

Table S4. Divergence times (and standard deviations in Ma) for Aizoaceae, Molluginaceae, Portulacineae and selected clades in Cactaceae, obtained from secondary Multidivtime dating analysis.

Taxon	age (stem)	age (crown)
Outgroups		
Aizoaceae	46.9 (3.4)	41.9 (4.0)
core Ruschioideae	33.0 (4.5)	17.1 (4.5)
core Ruschioideae excl. *Drosanthemum*/ *Chasmatophylum* group	17.1 (4.5)	7.0 (3.2)
Molluginaceae	53.3 (2.1)	46.2 (3.1)
Portulacineae	53.3 (2.1)	44.9 (2.7)
Montiaceae	44.9 (2.7)	39.9 (3.1)
Halophytaceae	42.5 (2.8)	
Didiereaceae *s.s.*	41.6 (2.8)	36.7 (3.8)
Basellaceae	39.8 (2.9)	29.4 (4.2)
Talinaceae	39.4 (2.7)	29.9 (3.3)
Anacampserotaceae	37.0 (2.6)	31.2 (3.0)
Portulacaceae	35.0 (2.6)	17.6 (3.3)
Cactaceae	35.0 (2.6)	28.6 (1.9)
Pereskioideae		
Pereskia (Northern)	28.6 (1.9)	25.8 (2.3)
Pereskia (Andean, SSA)	27.0 (1.7)	23.8 (2.3)
Opuntioideae	25.3 (1.2)	15.1 (2.9)
Opuntia	7.5 (2.3)	5.6 (1.9)
Maihuenioideae	24.4 (1.0)	4.8 (2.4)
Cactoideae	24.4 (1.0)	21.8 (1.7)
Blossfeldieae	21.8 (1.7)	
Cacteae	19.7 (2.0)	15.6 (2.4)
Mammillaria	6.3 (2.0)	3.7 (1.7)
Coryphantha	6.3 (2.0)	1.4 (1.1)
Turbinicarpus	6.6 (2.4)	2.1 (1.6)
Phyllocacteae	16.1 (2.3)	13.5 (2.5)
Echinocereinae	6.7 (2.0)	5.5 (1.8)
Echinocereus	4.6 (1.7)	2.8 (1.3)
Rhipsalideae	16.0 (2.2)	11.2 (2.7)
Rhipsalis	10.0 (2.5)	5.7 (2.5)
Notocacteae	14.8 (2.2)	12.1 (2.5)
Cereeae	14.8 (2.2)	11.8 (2.4)
Trichocereinae	6.5 (2.0)	5.3 (1.9)

Table S5. Species coverage in diversification analyses. Groupings in bold are paraphyletic.

Lineage	No. of species	percent coverage
PORTULACINEAE		
Cactaceae		
Pereskioideae	17	100
Mahuenioideae	2	100
Opuntioideae	349	83
Cylindropuntieae	123	71
Opuntieae	212	96
incertae sedis	14	
Pterocactus	9	0
Maihueniopsis (part)	4	0
Cumulopuntia (part)	1	0
Cactoideae	1498	71
Blossfeldieae	1	100
Cacteae	396	86
Phyllocactinae	308	56
Corryocactinae	75	58
Echinocereinae	146	51
Hylocereinae	87	61
Rhipsalideae	54	100
Notocacteae	104	98
Cereeae	589	61
Rebutiinae	150	78
Cereinae	161	32
Trichocereinae	278	69
incertae sedis	46	
Calymmanthium	1	100
Copiapoa	27	100
Frailea	18	100
Anacampserotaceae	36	100
Basellaceae	19	95
Didiereaceae	20	100
Halophytaceae	1	100
Montiaceae	226	71
Portulacaceae	116	100
Talinaceae	28	96
OUTGROUPS*		
Aizoaceae	1829	
Aizooideae	52	63
Mesembryanthemoideae	100	4
Ruschioideae	1584	47
Sesuvioideae	36	97
Tetragonoideae	57	0
Amaranthaceae	2050-2500	4
Caryophyllaceae	2200	29
Gisekiaceae	5	100
Limeaceae	23	91
Molluginaceae	87	93
Nyctaginaceae	395	2
Phytolaccaceae	65	43
Sarcobataceae	2	100

***due to poor sampling in outgroups we did not include them in diversification analyses**

Chapter 4.

Tackling the Molecular Dating Paradox: Underestimated Pitfalls and Best Strategies when Fossils are Scarce

Anita Lendel[1], Alexandre Antonelli[2], Reto Nyffeler[*1]

[1] University of Zurich, Institute of Systematic Botany, Zollikerstrasse 107, 8008 Zürich, Switzerland
[2] University of Gothenburg, Department of Biological and Environmental Sciences, Carl Skottsbergs gata 22B, SE-41319 Göteborg, Sweden
[3] Gothenburg Botanical Garden, Carl Skottsbergs gata 22A, SE-41319 Göteborg, Sweden

Corresponding author: Anita Lendel, Institute of Systematic Botany University of Zurich, Zollikerstrasse 107, 8008 Zürich, Switzerland; E-mail: anita.lendel@systbot.uzh.ch, alexandre.antonelli@bioenv.gu.se, reto.nyffeler@systbot.uzh.ch

Manuscript for Systematic Biology

Abstract

The estimation of absolute divergence times on the basis of molecular sequence data and information from the fossil record for rate calibration poses particular challenges when ages are estimated for clades that have a scarce or absent fossil record. We term this the "molecular dating paradox": taxa that lack a useful fossil record and whose divergence ages would be the most interesting to unveil are at the same time most sensitive to biases that cannot be assessed by evidence from fossils. We thoroughly evaluate the performance of three commonly adopted approaches to circumvent this situation: 1) "Distant Fossil Calibration approach", where the molecular data set for the focal group is expanded to include relatives with a sufficiently large sample of informative fossils; 2) "Study Group Placeholder approach", whereby age estimates for the focal group are derived from a higher level dating analysis that includes only few representatives of the focal group as placeholders; and 3) a "Secondary Calibration approach", which uses estimates derived from a previous dating analysis, rather than fossils, for time calibration. We evaluate these approaches empirically by estimating the age of cacti (family Cactaceae), a morphologically well-characterized clade of plants with high ecological, societal and evolutionary values, but which lacks an adequate fossil record for molecular dating. We investigate the trade-off between including numerous representatives of the order Caryophyllales as its closely related but fossil poor external group as well as more distantly related, fossil-richer lineages of eudicots. We perform a series of analyses based on 18 data sets differing in the number of taxa and sampling strategy for taxon and fossil selection, and compare the results derived from five different dating methods: uncorrelated lognormal relaxed molecular clock (UCLN, implemented in BEAST), nonparametric rate smoothing (NPRS, implemented in r8s), autocorrelated semiparametric relaxed clock (penalized likelihood [PL], implemented in r8s), using both branch-pruning (PLBP) and fossil-based (PLFB) cross-validation, as well as different strategies for obtaining optimal smoothing values, and a local molecular clock method as implemented in PATHd8. We evaluate the effect of taxon sampling density over the entire tree and/or of removing the upper age constraint bious fossil constraints located closely to the estimated nodes, as well as using multiple versus a single age constraint. We find that all methods behave relatively consistently when age constraints are abundant and well spread. But, age estimates are much more divergent and with a wider range of variance when nodes to assess are distant from the fossil constrained area of the tree. Our results suggest that in cases when fossils are scarce, dense and balanced taxonomic sampling – representing the diversity of the focal group as well as closely and distantly related taxa– appears to be the most adequate approach. However, the current trend to use more sophisticated relaxed-clock methods such as UCLN and PL is very much dependent and influenced by the sampling approach and the taxon sampling density applied. By showing that studies using a single molecular dating method on the basis of a single data set do not investigate the full range of variance of age estimations, we accentuate the importance of carefully assessing the robustness of results. We yield for Cactaceae a crown group age of

38.0 Ma (95% HPD 27.4 - 48.7 Ma) for the comprehensive data set with dense taxon and fossil sampling and applying the distant fossil calibration approach and using UCLN implemented in BEAST. Most other estimates fall into the range of the highest posterior density for this estimated age of cacti (mean 43.2 ± 12.7 Ma).

Keywords: molecular dating method, scarce fossil record, fossil calibration, taxon sampling density, distant fossil calibration, secondary calibration, Cactaceae

Introduction

Many comparative evolutionary studies rely on estimates of divergence times inferred from a molecular phylogeny calibrated against the fossil record or other means of calibration (Kumar, et al. 2005, Buschiazzo, et al. 2012). These techniques typically involve either local molecular clock analyses (e.g. PATHd8; Britton, et al. 2007) or different types of relaxed clock. The latter either apply non-parametric rate smoothing (NPRS; Sanderson 1997), semiparametric rate smoothing on the basis of penalized likelihood (PL; Sanderson 2002), or a relaxed molecular clock approach using Bayesian methods (e.g., Drummond, et al. 2006). Rates of molecular substitutions are either modeled as autocorrelated (i.e., for NPRS and PL) or uncorrelated (i.e., for UCLN).

The results of dating analyses are affected by numerous factors, including (1) analytical protocols for estimating sequence rate variation (e.g., Welch and Bromham 2005, Rutschmann 2006, Goodall-Copestake, et al. 2009), (2) density and choice of taxonomic sampling (Linder, et al. 2005, Hug and Roger 2007), (3) molecular sequence marker used (Bell, et al. 2005, Goodall-Copestake, et al. 2009), as well as (4) presence and distribution of long branches in the phylogeny (Magallón 2010). However, the calibration of relative time estimates by applying one or several fossil constraints, with associated assumptions and uncertainties is by far the biggest challenge (Graur and Martin 2004, Van Tuinen and Hadly 2004, Heads 2005, Taylor and Berbee 2006, Hug and Roger 2007, Ho and Phillips 2009, Sauquet, et al. 2012). The use of fossils as either fixed, minimum, or maximum age constraints, as well as their number and position on the phylogeny, either placed within (i.e., "internal calibration") or outside (i.e., "external calibration") of the focal group, are critical aspects to consider in molecular dating analyses (Lee 1999, Conti, et al. 2002, Magallón 2004, Reisz and Muller 2004, Rutschmann, et al. 2007, Ronquist, et al. 2012a, Sauquet, et al. 2012). Ideally, age estimation studies are based on multiple, probability-based distributed calibration nodes (Ho 2007) and the fossils are reliably dated, plentiful and well distributed among the lineages of the focal group. However, fossilization rates are in general exceedingly low and highly biased taxonomically, geographically, and temporarily (Smith 1994, Benton 2000). Indeed, for most taxa there may be intrinsic (e.g., soft body) or extrinsic (e.g., environmental) reasons for a poor fossil record.

For poorly fossilized groups, investigators may either apply a geological calibration, such as the formation of an oceanic island, or the emergence of a land bridge or a suitable life zone (e.g., Baldwin and Sanderson 1998, Heads 2011); or include one or several calibration points outside the main taxonomic group of interest (e.g., Near, et al. 2011). These approaches are generally considered sub-optimal when compared to direct fossil calibration within the group of interest, and have consequently met significant criticism (e.g., Heads 2005).

This situation results in what we would like to term "the molecular dating paradox": taxa that lack a useful fossil record and whose divergence ages would therefore be the most

interesting to unveil are at the same time most sensitive to biases that cannot be assessed by evidence from fossils. Relatively few studies have attempted to quantify potential biases on age estimations under this situation, deriving from the selection of methodological approaches and individual parameters, as well as from the application of taxon and fossil sampling strategies. In this study we compare three main approaches (Fig. 1) for estimating divergence times in taxa lacking a reliable and abundant fossil record, regardless of dating software:

1) "Distant Fossil Calibration (DFC) approach": the molecular data set for the focal group is expanded to include outgroup representatives that comprise a sufficiently large sample of applicable fossils.
2) "Study Group Placeholder (FGP) approach": age estimates for the focal group are derived from a higher level dating analysis that includes only few (minimally two) representatives of the focal group as placeholders.
3) "Secondary Calibration (SC) approach": age constraints derived from previous dating analyses are applied, rather than direct evidence from fossils.

For the DFC approach, the tree topology is relatively balanced, preferably proportionally to the diversity of the clade, as the external groups as well as the focal group are represented with multiple taxa. In contrast, in particular the FGP approach leads to highly sq taxon sampling. Often, the FGP and SC approaches are combined to yield age estimates for subclades of focal groups with a poor fossil record (e.g., Arakaki, et al. 2011, Fior, et al. 2013).

Based on a large phylogeny calibrated with multiple fossils from deeply attaching lineages as well as a fixed calibration point based on the first appearance of tricolpate pollen in the fossil record, Milne (2009) presented a possible detailed protocol that can be applied to dating analyses of all families and genera of eudicots, regardless of the incompleteness of their fossil record. For many empirical studies, relatives closest to the focal group lacking fossils for calibration also have a poor fossil record; we refer to this collection of closely related outgroup taxa as the "Closest Fossil-Poor external group" (CFP group). In such cases it may be necessary to expand the taxon sampling to include a richer and well-established set of fossils, by sampling further and more distant relatives in the molecular data set to be analyzed; we refer to these latter taxa as the "Distant Fossil-Rich external group" (DFR group).

Hugall and Foster (2007) find that the combination of saturation-driven compression of cladistically basal nodes and the use of phylogenetically deep calibration constraints, external to the focal group, will result in relatively older estimates of divergence times for the focal group. Ho et al. (2008) questioned, however, the observation of slower rates when deeper calibration points are used. We further elaborate and detail on the protocol of Milne (2009) by using multiple, both deep and shallow (in time and phylogenetic placement) outgroup fossil constraints that are distant to the focal group. The inclusion of the two types of external groups (i.e., CFP and DFR groups) with the focal group in one and the same age estimation analysis avoids the potential biases resulting from applying a SC approach, in

particular when estimated nodes are derived from a FGP approach. Analyses of large data sets, however, pose particular challenges to many of the different analytical methods currently used for age estimation. We develop here a comparative empirical study to explore different factors applying to such a DFC approach.

The Focal Group: Cactaceae

The temporal diversification of cacti has been assessed quite differently in the past, with age estimates ranging from more than 100 Ma to about 30 Ma (i.e., early Cretaceous to late Miocene) (e.g., Mauseth 1990, Hershkovitz and Zimmer 1997). Hence, the family Cactaceae, which is devoid of a reliable fossil record, is a well-suited focal group for a comparative study of different analytical approaches for age estimation. A putative cactus fossil from the Eocene (Chaney 1944) was shown to be wrongly identified (Becker 1962). Furthermore, persistent parts of cacti (i.e., spines, seeds) from packrat middens from the past several 100'000 years do not help in calibrating the deeper nodes of the cactus phylogeny (e.g., Chamberland 1997). For the same reasons we are reluctant to rely on a recent report of fossil pollen of *Pereskia* from the Miocene-Pliocene boundary (i.e., 6-7 Ma; Graham, et al. 2001). Indeed, there are only a few reports of fossils for lineages of the plant order Caryophyllales (see Supplement S1).

The cactus family comprises some 1,850 species (Nyffeler and Eggli 2010b), which are almost exclusively endemic to the American continent including the Caribbean islands. Cacti are prominent and characteristic plants of many semi-arid and arid regions (Gentry 1982, Prance 1994). They are, with the exception of the shrubby and broad-leaved species of *Pereskia*, morphologically very distinctive stem succulents (Barthlott and Hunt 1993).

Molecular phylogenetic studies of the past fifteen years revealed that the family Cactaceae is closely related to morphologically distinct lineages of Portulacaceae (Hershkovitz and Zimmer 1997, Applequist and Wallace 2001, Nyffeler 2007) within the order Caryophyllales. A revised family classification (The Angiosperm Phylogeny 2009, Nyffeler and Eggli 2010a) places Cactaceae in suborder Portulacineae together with families such as Anacampserotaceae, Didiereaceae, Portulacaceae s.str., and Talinaceae. The sister-group of Cactaceae remains unclear with competing hypotheses derived from analyses of different molecular sequence datasets (Applequist and Wallace 2001, Edwards, et al. 2005, Nyffeler 2007), however, rather conserved genetic markers from the plastid genome favor Anacampserotaceae as being sister to Cactaceae (Nyffeler 2007, Nyffeler and Eggli 2010a).

The origin of Cactaceae has puzzled researchers for centuries (e.g., Hunt and Taylor 1990, Hershkovitz and Zimmer 1997) and has led to speculations about their age of origin (Supplement S2). Traditionally, rather old ages have been invoked for the family in order to allow for the evolution of their distinctive features and wide distribution. For instance, the disjunct distribution area of *Rhipsalis baccifera* in South America and Africa (e.g., Barthlott 1983), the only species to naturally occur outside the Americas, has been proposed to date

back to the formation of the Atlantic Ocean some 110–90 Ma (e.g., Gibson and Nobel 1986, Mauseth 1990, Butterworth and Edwards 2008: "deep in the Cretaceous"). Recently, sequence divergence as estimated from molecular phylogenetic analyses of cacti and relatives revealed limited genetic variation, which led to much younger age estimations (Hershkovitz and Zimmer 1997, Landrum 2002). Broad-scale dating studies of angiosperm families suggested a crown group age for Cactaceae in the Miocene, either 11-18 Ma (Wikstrom, et al. 2001) or 21-22 Ma (Bell, et al. 2010). Based on very few samples of Cactaceae, and using the age of the Hawaiian islands to constrain the age of some Hawaiian *Portulaca* species (Portulacaceae), Ocampo and Columbus (2010) estimated the crown-group age of Cactaceae to be 10 (3.1 – 19.1) Ma. More recently, the age of cacti and major subclades were estimated by Arakaki et al. (2011). Applying a "two-step dating approach" on a large set of Portulacineae representatives and with a calibration based on two secondary age estimates derived from a study of eudicot representatives, Arakaki et al. (2011) reported a mean age of 35.0 (± 2.6) Ma, for the stem lineage age and 28.6 (± 1.9) Ma for the crown group age of Cactaceae.

Aims of this Study

The primary goal of this study is to investigate the range of age estimates for the temporal origin of the family Cactaceae (i.e., its crown group [CG] and stem lineage [SL] ages) and its major subclades depending on the applied sampling strategy and analytical method. Given the absence of a fossil record for cacti and their closest relatives, we investigate how the three different analytical approaches for a focal group devoid of a fossil record (i.e., DFC, FGP, SC) affect age estimates depending on the set of different samples of species and fossils. We expand our taxonomic sampling to include all major clades of Caryophyllales (our CFP external group) as well as many eudicot representatives (our DFR external group). For each approach, we employ and compare the following dating methods: uncorrelated lognormal relaxed molecular clock (UCLN, implemented in BEAST [(Drummond and Rambaut 2007)]); nonparametric rate smoothing (NPRS, implemented in r8s version 1.71 [(Sanderson 2003, Sanderson 2006)]); autocorrelated semiparametric relaxed clock (penalized likelihood [PL], implemented in r8s), using either branch-pruning (PLBP; Sanderson 2002) or fossil-based (PLFB; Near and Sanderson 2004) cross-validation, and a local molecular clock method as implemented in PATHd8 (Britton, et al. 2007).

Materials and Methods

Taxon Sampling and Sequence Matrix

We included a total of 460 species in our "exhaustive data set", of which 170 (c. 1.5% of a total of some 11,200 spp.) are Caryophyllales (excluding Cactaceae) and 129 (c. 7% of a total of some 1,850 spp.) are Cactaceae (Supplement S3). The selection of eudicot representatives was largely guided by the studies of Anderson et al. (2005) and Magallón and Castillo (2009). Three species of *Ceratophyllum* were included as outgroups. We assembled a data matrix of *matK* sequences by downloading 422 sequences from GenBank (www.ncbi.nlm.nih.gov/genbank/) and contributed 38 sequences from our own investigations (Lendel et al., submitted; Supplement S3). Sequences were aligned in MacClade 4.0 (Maddison and Maddison 2000) and indels were coded using the IndelCoder function of the software SeqState (Müller 2005) following the "simple indel coding" method (Simmons and Ochoterena 2000).

Phylogenetic analyses

We conducted a maximum parsimony (MP) analysis using PAUP* v. 4.0b10 (Swofford 2002) and a Bayesian phylogenetic analysis in MrBayes v. 3.1.2. (Huelsenbeck and Ronquist 2001, Ronquist and Huelsenbeck 2003) on matrix containing 460 *matK* sequences. Coded gaps were only included in the MP analysis. Most parsimonious trees were retained from a heuristic search consisting of 1000 random addition sequence replicates and the tree bisection-reconnection (TBR) branch-swapping algorithm. Relative support for all nodes was estimated using the bootstrap re-sampling procedure (Felsenstein 1985) implemented in PAUP*, using 1000 replicates, each with 100 random additions.

For the Bayesian analysis, the best-fitting model was determined with the Akaike Information Criterion using Modeltest v. 3.7 (Posada and Crandall 1998) and the resulting model parameters (TVM+I+G) were used as settings for the Bayesian analysis. Four independent parallel runs of four Metropolis-coupled Monte Carlo Markov Chains (MCMCMC) were conducted; each using a random tree as a starting point, running for $35x10^6$ generations, and sampling trees every 1000^{th} generation. The posterior output of MrBayes was examined with Tracer v. 1.5 (tree.bio.ed.ac.uk/software/tracer/) to assess mixing and convergence of MCMC chains, as well as to determine a suitable burn-in threshold. From the posterior trees one phylogram was randomly selected and used for the comparative study of the performance of the four different age estimation methods considering the three different analytical approaches outlined above. The selected topology was first compared with the strict consensus of the maximum parsimony analyses as well as with inferred relationships as published in several published studies (i.e., Bremer, et al.

2002, Cuénoud, et al. 2002, Nyffeler 2002, Edwards, et al. 2005, Soltis, et al. 2007) to check for any major incongruences.

Age Constraints

We selected 32 fossils distributed among the major lineages of eudicots, which we used to apply age constraints for our tree calibration (Fig. 2, Supplement S1). The eudicot crown group clade was constrained for almost all analyses, following the arguments provided by Milne (2009), with a fixed age of 125 Ma (Magallón and Castillo 2009; but see, Magallón 2010, Smith, et al. 2010, Doyle 2012). This calibration corresponds to the oldest records of tricolpate pollen (Forest 2009). Seven minimum age constraints were used for the early diverging eudicots, 19 for the core eudicots excluding Caryophyllales, and five for Caryophyllales, of which three are part of core Caryophyllales (i.e., Amaranthaceae, Caryophyllaceae, and Phytolaccaceae). The estimated age of these fossils and their assignment to either SL or CG position follow, largely, Anderson et al. (2005) and Magallón and Castillo (2009) (see Supplement S1 for detailed information on fossils and references). Overall, nodes were constrained according to the method of "upper limit" trial (Hug and Roger 2007) with the fossil age as a minimum age boundary for the node, and a global maximum age limit (in our case a fixed age of 125 Ma) for all constrained nodes.

Data Sets for the Comparative Study

The exhaustive data set (hereafter data set “dsA”), comprising dense and comprehensive taxon and fossil sampling (460 *matK* sequences, 32 fossil constraints) as well as the randomly selected phylogram from the Bayesian analysis formed the basis from which 17 alternative data sets were compiled (see Table 1). These differ (1) in the total number of representatives, (2) in the relative number of species and fossil representatives (either from the DFR external group or from the focal group), (3) in whether or not a fixed constraint for the eudicot crown group clade was enforced, (4) in whether or not rather controversial fossil constraints from the core Caryophyllales were included in the analyses, or (5) in only applying a single, fixed fossil constraint. For the data sets with a reduced taxon sampling the selected phylogram was pruned from the terminal branches and branch lengths were newly estimated from the sequence data and the selected model using PAUP* v. 4.0b10. Datasets B, C, and D (dsB, dsC, dsD; Table 1) represent the DFC approach with either 345, 230, or 115 taxa (corresponding to ¾, ½ and ¼ of dsA; see Supplements S4a to S4c, S5). Data sets E and F (dsE, dsF; Table 1) represent the FGP approach and G and H (i.e., dsG, dsH; Table 1, Supplement S5) represent the SC approach, each derived from either dsA or dsD, respectively. In addition, dsI and dsJ, derived from dsA and dsD as well, do not contain a fixed maximum age constraint, and dsK and dsL do not consider controversial fossils from the CFP external group. Finally, dsM, dsN, dsO, dsP, dsQ, and

dsR, derived from dsA, dsD, dsE, dsF, dsG, and dsH, only apply a single, fixed calibration of 125 Ma for the crown group of eudicots.

Molecular Dating

Tests for molecular clock violation. - For all data sets and corresponding phylograms we tested for the violation of clock-like evolution by performing likelihood ratio tests (Felsenstein 1981). Additionally, all phylograms were analyzed with the clock test as implemented in the software PATHd8.

Age estimation analyses - We estimated divergence times for eleven selected nodes based on the phylograms for dsA to dsR with four relaxed clock methods that implement among-lineage rate heterogeneity differently: (1) a Bayesian relaxed-clock method with an uncorrelated lognormal (UCLN) model implemented in the software BEAST, (2) an autocorrelated (Gillespie 1991) nonparametric relaxed clock (nonparametric rate smoothing, NPRS), (3) a semiparametric relaxed clock (penalized likelihood [PL] with two different rate smoothing optimization criteria), and (4) a nonparametric method that optimizes mean substitution rates locally in segments of the topology (PATHd8).

Uncorrelated lognormal dating. This method was used for all data sets to estimate divergence times but also phylogenetic relationships simultaneously. The topology of the selected phylograms was used as starting tree for the BEAST analyses. The model applied to each data set was GTR+I+G, accommodating for among-site rate variation and proportions of invariable sites, with 4 gamma categories and a Yule prior for the different analyses. Several taxon sets were specified with the MRCA command, and the eudicot crown group was enforced to be monophyletic. The prior age of the root was modeled with a mean of 125 Ma for a normal distribution with a standard deviation of 2.0. All other fossil constraints, except for two fossils detailed below, were modeled with a lognormal distribution, where 95% of the prior distribution is restricted to the geological period to which it was assigned. However, the age constraints for the SL for Phytolaccaceae and the SL for Vahliaceae were assigned priors with a normal distribution with mean fossil age and a standard deviation of 2.0. BEAST analyses were run on a personal computer (Processor: Intel(R) Core Dual Processor E7500) and at the Cornell Computational Biology Unit: cbsuapps.tc.cornell.edu/beast.aspx. Tracer v. 1.5 was used to visualize performance of each chain and assess convergence of chains and effective sample sizes (ESS values) for all relevant parameters.

We carried out an iterative process of BEAST analyses, testing results and setting up additional runs. Up to $30x10^6$ generations were thus collected per data set, sampling every 1000^{th} generation. A burn-in cut-off of 10% was generally applied, unless a larger cut-off was needed based on observed chain performance. The post burn-in trees from all runs were re-sampled at lower densities to reduce file sizes and combined using LogCombiner v1.5.3

(beast.bio.ed.ac.uk/LogCombiner). Maximum credibility trees were then calculated from more than 10'000 trees or more for each data set, using TreeAnnotator v1.7.4 (beast.bio.ed.ac.uk/TreeAnnotator). The mean and 95% highest posterior density (HPD) for the divergence time estimates of the selected clades were inspected on the chronograms using FigTree v.1.3.1. (tree.bio.ed.ac.uk/software/FigTree/). Additionally, we performed a MCMC run without any sequence data in order to investigate the relative influence of the data and the prior settings on the results (Drummond, et al. 2006).

NPRS dating. - We estimated divergence times with the NPRS method as implemented in the software r8s for selected clades based on the selected phylogram that was modified by pruning taxa and newly calculating branch lengths to fit dsA to dsR. This method optimizes rates simultaneously over the entire phylogram while minimizing ancestor-descendant local rate changes under the assumption of autocorrelated rate changes. It allows fast analyses of large phylograms and was successfully applied to all data sets with the fixed age of 125 Ma for the eudicot crown group included.

Penalized likelihood and cross-validation procedures. - Divergence times for selected clades were estimated with the autocorrelated, semiparametric dating method PL using the software r8s (Sanderson 2006) on the phylograms of the dsA to dsR. This method assumes divergence rates among lineages to be autocorrelated, but reduces their fluctuations by a smoothing parameter chosen in a data driven cross-validation procedure. PL offers two different approaches to calculate the optimal rate smoothing parameter (λ). We adopt the abbreviations used by Magallón (2010) for these two types of cross-validation procedures: PLBP for penalized likelihood with branch-pruning cross-validation and PLFB for penalized likelihood with fossil-based model cross-validation. The commonly used PLBP cross-validation method relies on pruning successively the terminal branches from the tree and estimating, based on the information on the estimated number of substitutions on each branch, the smoothing parameter that corresponds to the least prediction error (Sanderson 2002, Near and Sanderson 2004). An alternative to this is the cross-validation procedure that relies on the information provided by fossils rather than molecular data, PLFB. In this procedure, minimum and/or maximum age constraints on nodes are successively removed, times and rates across the tree are re-estimated, and, if present, a violation of the fossil-derived age is determined and scored. An overall cross-validation score is summed across the tree and calculated for a range of smoothing parameters, resulting in an optimal smoothing value that corresponds to the lowest cross-validation score (Near and Sanderson 2004). For each of the two cross-validation procedures we tested a wide range of smoothing values ranging from $\log_{10}\lambda = -5.00$ to $\log_{10}\lambda = 9.00$, in rather large (i.e., 0.5 and 0.25) and small (i.e., 0.10 and 0.01) increments. PL analyses used the Truncated Newton algorithm, 5,000 maximum iterations, and the check gradient option enforced with 5 restarts and 5 guesses. The outgroup taxa, either *Ceratophyllum* for all data sets except dsG and dsH, or *Dillenia* for ds G and dsH, were pruned prior to cross-validation analyses and zero-length

branches were collapsed. For each performed cross-validation, the smoothing parameter chosen as optimal was the one that corresponded to the least predicted error for a given data set.

Additionally, we examined the variation of the optimal smoothing parameters depending on the range of smoothing values and the size of the selected increments in the cross-validation procedure. Often, in order to speed up the calculation of the optimal smoothing parameters, an analysis is run first with large increments and consecutively the range is narrowed and the increments is set to smaller values. For instance, the cross-validation procedure is started at $\log_{10}\lambda$ = -5.00 with an increment of 0.5 for a certain number of smoothing magnitudes. In a second step, only the range for which the smoothing parameter has passed the cross-validation analysis is further tested with a smaller increment, and the smoothing parameter with the lowest chi squared error is chosen as the optimal.

The "best" optimal smoothing parameter for each of our data sets was selected based on testing the entire range of smoothing values in a single step. Despite repeated efforts and regardless of the type of cross-validation applied, the software was not successful in analyzing data sets with more than 115 taxa and more than one fossil constraint (except for dsG, with 300 taxa and 6 fossils constraints).

PATHd8. - Based on the assumption of a local molecular clock in tree segments, this dating method optimizes the mean substitution rates between sister taxa by sequentially taking averages over path lengths from an internode to all its descending terminals (Britton, et al. 2007). We applied this method to phylograms of all data sets, from dsA to dsR.

Results

The aligned *matK* sequence matrix of dsA comprised, after pruning part of the 5'- and 3'-ends, 1824 characters. Gap-coding contributed an additional 286 parsimony-informative characters. Altogether, 1453 characters were informative and 284 were constant.

The Bayesian analyses in four independent runs of 50 million generations were checked for convergence, and then the burn-in threshold was set to 10%. Hence, a total of 10'000 posterior trees were randomly selected from 135'152 trees to construct a 50% majority-rule consensus phylogram with mean branch lengths (Supplement S6). The first topology was selected for age estimation analyses using NPRS, PL, and PATHd8. The parsimony analysis yielded 750 trees of 19,236 steps, with CI=0.81 and RI=0.70 (Supplement S7).

The likelihood ratio test for rate-constancy and the mean path length (MPL) measure implemented in PATHd8 rejected, with very low p-values, the hypothesis for a constant molecular rate for the data sets. Finally, BEAST, implementing the uncorrelated lognormal (UCLN) model, estimated simultaneously topology and branch lengths from the aligned 460

matK sequences and 32 fossil constraints (Fig. 3). We compared these three topologies for the presence of more than 60 clades and sister-group relationships. Overall, inferred relationships were very similar (Supplement S9a and S9b) and congruent to the randomly selected phylogram from the posterior trees of the Bayesian analyses.

Impact of Taxon and Fossil Sampling on Dating

Age estimates for eleven nodes (SL and CG divergence times for Asterids as well as for Caryophyllales, Portulacineae, Cactaceae, Cactoideae, and Trichocereinae) derived from dsA to dsR are reported in Supplement S8. Fig. 4 depicts results for estimated divergence times for eleven nodes and twelve different data sets (dsA to dsL).

UCLN was the only dating method that provided estimates of divergence times for all 18 data sets. NPRS, PL, and PATHd8 were applicable for data sets without a fixed age constraint (dsI and dsJ). In addition, for PL, both types of cross-validation procedures (fossil-based or branch-pruning) failed to find optimal smoothing parameters for dsA, dsB, dsC, dsE and dsK, and, therefore, were not considered. Furthermore, PLBP was equally unsuccessful in estimating the optimal smoothing parameter of dsG, dsM, dsO and dsQ. Fossil-based model cross–validation requires at least one fixed and two constrained ages, and was therefore not performed for the data sets with fewer constrained ages (dsM, dsN, dsO, dsP, dsQ and dsR).

Performance of BEAST Analyses

The comparison of ages resulting from the analyses in which BEAST sampled only from the prior distribution and the runs with data revealed considerable differences, indicating that the data contained substantial information for the analyses performed here. The number of generations necessary to reach convergence and adequate ESS values (recommended to be higher than 100 – 200 for all relevant parameters) varied, depending on the data set. We computed some qualitative (e.g. mixing and convergence) and quantitative (e.g. the number of runs combined and the duration of each run) statistics to evaluate the performance of multiple BEAST analyses for each data set (summarized in the Supplement S8). For almost all of the data sets analyzed, MCMC chains from multiple independent runs showed good convergence and mixing and were therefore combined. Only in two cases, dsB and dsO convergence in a large number of runs did not reach the threshold and we do not consider those results further. In all other cases we combined multiple and occasionally numerous (up to 20) individual runs in order to reach ESS values adequate for parameters applied. In general, data sets with dense taxon sampling required two to three times longer to finish the analysis than their subsets with sparsely sampled taxa.

Finding the optimal smoothing parameter in PL

Penalized likelihood with two different smoothing optimization criteria was used to estimate divergence times for all data sets, but failed for several data sets with more than 115 taxa and, in most cases, more than 1 fossil constraint (see Supplement S8). Cross-validation for PL proved to be unreliable, strongly affected by the start and the end point of calculation, and particularly dependant on the value of the increment. The magnitude of the optimal rate smoothing value and the pattern of the cross-validation error versus $\log_{10}\lambda$ plot differed unequivocally for dsD between the two rate smoothing optimization criteria of PL, corroborating the observation of Magallón (2010); see supplements S10a and S10b. Furthermore, regardless of the cross-validation process applied, testing the entire range instead of limiting the range by several consecutive steps resulted, usually, in different optimal smoothing parameter values (Supplements S10a and S10b).

Discussion

Contrasting the DFC, FGP, and SC approaches

In general, when estimating ages with the three different sampling approaches outlined in this study (DFC, FGP, and SC), analytically more sophisticated methods like UCLN and PL yielded a higher variance of age estimates, while NPRS and PATHd8 produced rather similar estimates, independent of taxon density (Figs. 5a to 5h, Supplement S8). In comparison to estimates yielded by the DFC approach, the FGP approach regularly estimated younger ages, while the SC approach, to various degrees, yielded either younger (i.e., UCLN), older (i.e., PLBP and PATHd8) or similar age estimates (i.e., NPRS, PLFB). The 95% HPDs observed in the UCLN analyses were smaller in data sets with densely sampled taxa and in particularly for those with the DFC and the SC approach. Among the three approaches, larger differences were observed for node ages further away from the fossil constrained area of the tree, with the most prominent differences yielded by the UCLN analyses between the DFC and the FGP approach. For example, for UCLN the CG Cactaceae was estimated to 18.9 million years earlier in dsE compared to in dsA, and 45.6 million years earlier in dsF compared to in dsD. PLBP, however, estimated in both the DFC and the FGP approach almost identical ages for all nodes. This indicates that major underrepresentation of the taxon diversity of the focal group, by sampling only a minimal number of taxa and in particular when taxon sampling overall the tree is sparse, has a huge impact on estimating ages with UCLN, but less so with NPRS, PL, or PATHd8.

Interestingly, the largest differences in ages estimated with the SC approach, when compared to ages estimated with the DFC approach (i.e., dsH compared to dsD), were yielded by the PLBP, which considerably overestimated ages. The second largest difference, this time the SC approach underestimating ages, was observed between data sets with dense taxon sampling (i.e., dsA and dsG) analyzed with UCLN. In the SC approach PATHd8 somewhat overestimated ages for both densely and sparsely sampled data sets, which could be explained by constraining the age of the CG Caryophyllales to an age (i.e., 100 Ma) that is slightly older than the PATHd8 estimates for this node (i.e., 91 Ma) in respective data sets applying the DFC approach.

Confronted with a scarce fossil record in the focal group, all three methods outlined here have been widely reported in the literature. A number of authors performed dating analyses that could be classified as representing the DFC approach - using fossil calibrations distributed among more or less closely related outgroup/external species (Yang and Yoder 2003, Near and Benard 2004, Goodall-Copestake, et al. 2009, Milne 2009, Michalak, et al. 2010, Pitts, et al. 2010, Near, et al. 2011). Others have estimated ages using a large-scale phylogeny, with only few (sometimes a single) representatives of the focal group (i.e., representing the FGP approach(e.g., Wikstrom, et al. 2001, Moore, et al. 2010). Still others have performed age calibration deriving rates from previous dating analyses (i.e., SC approach; (e.g., Goldblatt, et al. 2002, Franzke, et al. 2009, Zhang and Fritsch 2010).

More often however, divergence times of a fossil scarce group were estimated in a two-step strategy, combining the FGP approach and SC approaches (e.g., Van Tuinen and Hedges 2004, Pfeil and Crisp 2008, Janssens, et al. 2009, Su and Saunders 2009, Arakaki, et al. 2011, Wilkinson, et al. 2011, Nylinder, et al. 2012). Initially, a broadly sampled phylogeny, with few placeholder taxa representing the focal group, is dated in order to estimate the divergence times of several nodes of interest. In a second step, divergence time(s) as estimated in the previous step are used as the starting point for additional dating analyses of the expanded focal group phylogeny. Often, reported ages for a given clade differ considerably from one study to another (Pulquerio and Nichols 2007) and we believe that those large differences might be the result of the sampling approach applied, besides different dating methods and fossil constraints. Indeed, applying either the DFC, the FGP, or the SC approach, or combining the latter two, might be part of an explanation for the large differences in the age estimates for Caryophyllales and Cactaceae as estimated in different studies (Supplement S2).

Age estimates for the major angiosperm clades in the current analyses

Following a seminal study estimating the temporal origin of major angiosperm clades by Wikström et al. (2001), Bell et al. (2010), Moore et al. (2010), and Smith et al. (2010) have estimated divergence times for major lineages of angiosperms relying on the UCLN model implemented in BEAST. Using different data sets, taxon sampling, and fossil calibrations, they represent differently the true diversity of Caryophyllales (ca 11 500 spp.;Buchheim 1964) and report considerably different ages for its CG node. Bell et al. (2010) samples 24 taxa from within Caryophyllales (out of 567 angiosperm lineages overall, calibrated with 36 fossil constraints), and infer their CG age to 99-106 (91-115) Ma. Smith et al. (2010) represent Caryophyllales with two taxa (out of 154 species of land plants, calibrated with 33 fossil constraints) and date their CG age to 83.5 Ma. Finally, based on data of 86 species of seed plants, constrained with 4 fossil calibrations, Moore et al. (2010) represent Caryophyllales with two taxa and estimate the CG age to 67 (63-71) Ma.

From the perspective of Caryophyllales as the focal group, the outlined studies could here be classified either as representing the DFC approach (Bell, et al. 2010) or the FGP approach (Moore, et al. 2010, Smith, et al. 2010). The younger age of Caryophyllales in the latter two studies, and particularly in the study of Moore et al. (2010), with a scarce taxon sampling, would then correspond to the pattern we recover in our study with the FGP approach, in comparison to the DFC approach. However, if Cactaceae is chosen as focal group, the same study of Bell et al. (2010), in which cacti are represented with a single taxon, can be classified as representing the FGP approach, and therefore comparable to our dsE, which contains two representatives for cacti. The estimated ages for those two studies are also comparable, while Bell et al. (2010) estimated SL Cactaceae to 21-22 (11-33) Ma and the UCLN analysis of data set dsE in this study yielded a mean age of 19.1 (0.8-43.4) Ma. Combining the FGP approach and the SC approach, Arakaki et al. (2011) estimate in

MultiDivTime (Thorne and Kishino 2002) SL Cactaceae to 35 (±2.6) Ma, while applying the DFC approach in UCLN we estimate an older age, 47.1 (36.2-57.8) Ma.

A similar pattern of ages being highly dependent on taxon sampling and methodology has been reported in numerous other studies. For instance, Wikström et al. (2001), estimated Orchidaceae to be 69-53 (±5) Ma (SL) and 26-37 (±3) Ma (CG) old, whereas Janssen and Bremer (2004) estimated them to be considerably older, 119 (SL) and 111 (CG) Ma, , while although both analyses used NPRS and a single (but different) fossil calibration. However, in the analyses of 560 species of angiosperms, Wikström et al. (2001) sampled only two genera of Orchidaceae (out of 880 genera recognized within Orchidaceae (Buchheim 1964), and therefore in respect to this family, employed an unbalanced sampling of the true diversity - corresponding to the FGP approach of Janssens and Bremer (2004), in the study of 878 genera representing all ten orders of monocots, sampled extensively Orchidaceae for 145 genera, approaching in respect to their focal group, to a more balanced sampling and therefore to the DFC approach. The pattern described in the latest example corresponds to a general underestimation of ages yielded in our study by the FGP in comparison to the DFC approach; exacerbated, also within NPRS, when a single calibration is applied (see below, Figs. 5a to 5h, Supplement S8).

Although recognized as a probable source of error (Hugall, et al. 2007), the impact of a sparse and unbalanced taxon sampling on divergence times has not received considerable attention (Soares and Schrago 2012). Estimating divergence times based on a poorly sampled focal group (i.e., FGP approach) may seem like a straight-forward solution, but may also lead to large estimation errors (Van Tuinen and Hedges 2001) and may, as reported in our study, underestimate age estimates.

Surprisingly, the often-criticized SC approach proved in our study to perform better, given that the starting age (i.e. secondary calibration point) corresponds to the most reliable estimate of the node's age, and that multiple calibration points are constrained. In the divergence time calibration study of the fossil-rich plant genus *Nothofagus*, Sauquet et al. (2012) tested the accuracy and the reliability of the secondary calibration approach by constraining core Fagales to ages derived from previous studies and from within their own study, to find that this approach yielded drastically younger ages. We observe, up to a certain point, the same pattern in our data sets with multiple (dsG and dsH) and especially with a single (dsQ and dsR) calibration constraints. However, when ages used for calibration in the SC approach already underestimate the age of a node, this underestimation will be even more accentuated and the younger ages will be estimated for the focal group. Therefore, we believe that combining the FGP the SC approaches should estimate younger ages, as shown for age estimates for cacti yielded in the study of Arakaki et al. (2011) compared to this investigation. As an alternative solution, we suggest that when fossils are scarce the DFC approach should be used to overcome the negative effect of the FGP approach.

Sauquet et al. (2012) tested with a well balanced ingroup and outgroup sampling (27 and 21 taxa, respectively) where all available outgroup and ingroup fossil constraints were

applied, and compared them to a scenario corresponding to our DFC approach in which they considered only outgroup constraints (scenarios 1 and 3, respectively in their study). They found that the different sets of constraints did not produce drastically different ages across the tree, supporting the DFC approach as an adequte alternative in estimating divergence ages for a group without suitable internal age constraints.

Our finding that the undersampling of a given clade can influence divergence time estimates differs from the general finding of Soares and Schrago (2010) suggesting that the effect of topological tree shape is negligible. However their study set-up differs from ours, and likewise some of their findings may not be directly comparable. Our estimation of divergence times in deep nodes (e.g., nodes D1 and D2 in Soares and Schrago (2012)) thus allows us to reassess how such nodes are influenced by an (un)balanced representation of the taxa in the clade itself. Soares and Schrago (2010) additionally detected that increasing the number of terminals between the calibration node and the shallow nodes improved the age estimate of the shallow divergences. We therefore do not support the recommendation of Soares and Schrago (2010, page 137) that in order to improve divergence time estimates, for a given set of calibrations "... it is better to add two sequences (nodes) on the path between the calibration node and the node of interest than it is to add one on this and the other on the sister group." Instead, in order to estimate rate differences well in the considered phylogeny and in line with the findings by Linder et al. (2005, page 574) who reported that "... only the average sampling density on the whole tree under investigation has an impact on the results, not the sampling density of the individual nodes". We therefore encourage a dense and balanced taxonomic sampling whenever possible, representing well the diversity within the focal group and to its closely and distantly related taxa.

The difficulties related to the DFC approach are more practical and applied, as sequences across a wide taxonomic range, in our case including representatives from all major lineages of eudicots, need to be sampled for one or, preferably a set of regions. Furthermore, the molecular dating methods capable of analysing large amounts of data are needed. The first issue can be avoided by applying a supermatrix approach, in which matrices of different regions overlapping in a number of taxa are combined. The issue of analyzing large data sets can today be overcome by using methods employing Bayesian relaxed molecular clock (e.g., programs BEAST, MultiDivTime, and the recently updated version of MrBayes 3.2 (Ronquist, et al. 2012b). Still, as eminent from this study, for all those molecular dating methods, future studies need to assess their sensitivity to taxon sampling the (un)balanced tree topologies.

Effects of Taxon Sampling Density

We were able to observe the effect of different taxon sampling densities over the entire phylogeny, for both external groups and Cactaceae, while constraining 32 fossils (i.e., DFC approach) in UCLN, NPRS and PATHd8 analyses (Figs. 6a to 6d). We were not able to

assess the sensitivity of PL to taxon sampling density as the cross-validation, regardless of type, failed to run for data sets dsA, dsB and dsC (Supplement S8). We found that the variance among the different data sets was higher with an analytically more sophisticated method such as UCLN than with simpler methods such as PATHd8 or NPRS. The range of ages in this comparison (dsA to dsD) found for CG Cactaceae using UCLN was 38.0 to 63.1 Ma (with HPD of 27.4 to 83.7 Ma), while for NPRS it was 40.5 to 50.6 Ma. For the nodes we evaluated in the area of the tree constrained by fossil ages (i.e. CG asterids, SL and CG Caryophyllales), the UCLN method estimated somewhat younger ages with the sparse taxon data set, a pattern that changed radically in the part of the tree without age constraints (i.e., SL and CG nodes of Portulacineae, Cactaceae, Cactoideae and Trichocereinae). Sampling taxa sparsely provided in the part of the tree without fossil constraints substantially older UCLN estimates with a wider HPD, a trend that increased considerably the further away the node was from the fossil constrained part of the tree (as revealed by the coefficient of variation in Figs. 6a to 6d). While PATHd8 estimated comparable ages regardless of taxon density, NPRS estimated somewhat younger ages when taxa were sampled sparsely, particularly in the area of the tree without age constraints and for nodes further away from the closest calibration point. The coefficient of variation did not reveal a correlation of the undersampling effect and the distance from the calibration node.

It is also interesting to observe the patterns revealed with the FGP and the SC approaches when taxa over the whole phylogeny were sampled with different densities (Figs. 6e to 6h). Contrary to the pattern found in the DFC approach, sparse taxon sampling in the FGP approach produced older ages of the nodes in the fossil constrained area; and younger ages of the nodes in the area of the tree without age constraints. This indicates that, when analyzed with UCLN, undersampling taxa has different effects on age estimates, depending on the sampling approach applied. Indeed, in the UCLN we found the biggest difference in the mean estimate for the CG Cactaceae between the sparsely sampled data set applying the DFC approach (dsD: 63.1 Ma) and the sparsely sampled data set applying the FGP approach (dsF 17.5 Ma). Applying the SC approach, we were able to analyze in the PLFB a single data set with a dense taxon sampling (i.e., dsG), which yielded older ages than when taxa were sampled sparsely. Once again differences revealed with NPRS and PATHd8 were much smaller, confirming the observation that more sophisticated methods (i.e., BEAST, PL) are more sensitive to taxon sampling density over the whole tree and particularly when different sampling approaches are applied.

Despite taxon sampling being recognized as a crucial issue not only in phylogenetic reconstruction but also in molecular dating (e.g., Heath, et al. 2008a), studies that rigorously examine and compare sensitivity of dating methods to taxon sampling density, while preserving all other conditions, are rare. A generally accepted premise is that dense taxon sampling pushes divergence time estimates back in time, especially in methods that apply smoothing between ancestor-descendant lineages, which will systematically result in overestimates (e.g., Janssen and Bremer 2004). Heath et al. (2008b) demonstrated in a simulation study that a random pruning of taxa from the trees simulated with variable and

autocorrelated speciation rates results in an increase in the average age of the internal nodes. Although increased taxon sampling will produce better branch lengths estimates and more stable results, it will also affect the results in different ways and different methods will not converge to similar age estimates (Anderson 2007). NPRS in general was found to be more sensitive to reduction in taxon sampling density than PL (Linder, et al. 2005, Pirie, et al. 2006, Milne 2009). Linder et al. (2005) evaluated in a nearly completely sampled clade of African Restionaceae, how increasing and decreasing of taxon sampling does effect node age estimates using NPRS, PL and the Bayesian methods. Similar to our study, they produced numerous smaller data sets by deleting terminals from the biggest data set, while preserving representatives of the basal lineages for each of the clades. They find that all methods are sensitive to taxon undersampling, and that age discrepancies are positively related to distance from the calibration node. Furthermore, they calibrated phylogenies with a secondary age estimate, obtained in their previous molecular study - it represents, therefore, a single secondary calibration approach (corresponding to our dsQ and dsR). Anderson (2007) suggests a large influence on the number of taxa for internal node ages under PL and MultiDivTime, and very little influence when estimated by PATHd8. The study of Jian et al. (2008) reveals that taxon density greatly affected ages of Saxifragales estimated in PL, however, the authors did not specify which cross-validation method was applied. Ages estimates based on a sparse data sets were in general substantially older than those estimated in the densely sampled data set.

The set-up of our study differs from Milne (2009), and therefore our results are not directly comparable with his "taxon sampling variation" - where up to three taxa outside the target group were added or removed from a full data set of 99 taxa; nor his "taxon sampling density" assessment - where large sets of taxa were added or removed from specific clades, instead from the whole tree as in our study. Milne (2009) concludes that all of those taxon sampling variations had small to negligible effects. It is interesting, however, to observe that deleting ten of the most closely related taxa to each target node in his "analysis D" had "dramatic" effects in estimating considerably older ages and confidence ranges in *Pachysandra* and *Podophyllum* – clades for which those deleted taxa represented the closest outgroups; confirming the importance of sampling the nearest outgroup (Robinson, et al. 1998). Milne (2009) also concludes that increasing taxon sampling in a clade sister to, or in a clade distant to the clade containing target node ("analyses dsF and dsE", respectively, in his study) had small or negligible effects to the initial estimates. This observation provides further support that our DFC approach will not alter the ages estimated for the focal group, but will allow, in our opinion, wider choice of calibration points.

Xiang et al. (2011) estimated in BEAST divergence times in the plant order Cornales and found that reducing taxon sampling density produced congruent estimation of divergence times; still the age of the crown node Hydrostachyaceae was estimated to be younger in analyses with reduced number of taxa, a difference they attributed to a random pruning of deeper diverging species. Interestingly, in their total evidence data set, they sample extensively outgroups from the Asterids clade, use multiple basal nodes for

calibration and include representative taxa from all Cornelean genera and families, with an extensive sampling of Hydrostachyaceae – an approach that could, therefore, be classified as our DFC approach. Additionally, fossils they use for calibration were constrained to Cornaceae and Nyssaceae - clades that are more distantly related to Hydrostachyaceae. In the reduced data set, however, the crown node of Hydrostachyaceae was represented by two species – therefore we believe that the differences they report in estimated ages for the clade in question are result of the underrepresentation of that clade in the smaller data set (i.e., FGP approach), which can easily have as a consequence pruning of the deep branching species. This observation only confirms their statement that the crown age of a clade in a phylogeny should be cautiously interpreted. In our study, analyzing in BEAST data sets with different taxon sampling density, in the part of the tree constrained by fossil ages, also produced relatively congruent estimation of divergence times and, therefore, congruent findings with Xiang et al. (2011). In general, we found numerous studies that analyze the "taxon sampling density effect" but only change the density of a given clade and not the density across the entire tree (e.g., Yoder and Yang 2004), confounding the effect that taxon sampling density has on age estimates.

Estimating Divergence Times with Different Relaxed-Clock Methods

Under the DFC approach with exhaustive taxon and fossil sampling (i.e., dsA) we were able to compare the ages estimated with three different dating methods. Sampling sparsely one fourth of that taxon sampling (i.e., dsD) in addition allowed estimation of ages by all dating methods investigated in this study, i.e. UCLN, NPRS, PL (with both cross-validation types) and PATHd8 (Figs. 7a and 7b, Supplement S8). All methods behaved relatively consistently when age constraints were abundant and well spread – as in the fossil-constrained area of our tree – but became more variable when close constrained nodes were absent. Ages varied substantially, more in nodes distant from the fossil-constrained area of the tree. For dsA, NPRS frequently provided older estimates (e.g., 50.6 Ma for CG Cactaceae); however those ages mainly coincided with the range of means (and HPDs) of the UCLN. PATHd8 yielded the youngest ages for all nodes, regardless of taxon sampling density. For the nodes we investigate in the fossil constrained area of the tree, the oldest age estimates were yielded by PLBP and the youngest (not considering the PATHd8 analyses) by the UCLN analysis of dsD. For the nodes in the area of the tree without age constraints, the observed pattern was the opposite and the UCLN analysis of dsD yielded the oldest ages (e.g. 63.1 Ma with an HPD of 41.5 to 83.7 for the CG Cactaceae) while the youngest ages (not considering PATHd8 analyses) were estimated by the PLBP (e.g. 26.4 Ma for CG Cactaceae).

The uncorrelated lognormal (UCLN) relaxed-clock model of Drummond et al. (2006) is today the most popular and most widely used dating method, as it allows uncorrelated rates of substitutions across the tree and incorporates uncertainties in tree topology and multiple, differently shaped fossil calibration priors. Its methodology and assumptions are

often readily compared to other dating methods in terms of analytical approaches, but this is not the case with their empirical outcomes. Comparisons of ages estimated by UCLN with those estimated by other dating methods are scarce due to operating in different ways (i.e., requirements for fossil calibration codings) or due to not analyzing identical data sets. In many studies UCLN yielded older estimates than PL and NPRS (Pitts, et al. 2010), or PL alone (e.g. Nie, et al. 2008, Goodall-Copestake, et al. 2009, Gustafsson, et al. 2010, Zhang and Fritsch 2010, Sauquet, et al. 2012), while in others estimates yielded younger ages than those estimated by PL (Su and Saunders 2009). When differences were detected, many of these studies (except Goodall-Copestake, et al. 2009, Pitts, et al. 2010) favored ages estimated under UCLN. We find in our study similar trends to those previously reported, which suggests that our focal group is not essentially different from others. Nevertheless, we detected substantial differences within this focal group depending on the data set used, particularly regarding taxon sampling density. The patterns found are striking and complex. For instance, comparing UCLN with other dating methods for the nodes in the area of the tree without age constraints (Supplement S8) revealed that UCLN estimated younger ages than NPRS but older ages than PATHd8 for dsA (Fig. 6a) , whereas UCLN yielded the oldest ages for dsD (Fig. 6b).

In general, NPRS was shown to overfit the data and in certain cases lead to an overestimation of divergence times (Sanderson 2002), a result supported by a number of subsequent studies and now regarded as a common point of view. For instance, a number of studies found that NPRS estimated older ages than the global constant clock, PL (PLBP/Ref.) a Bayesian dating method (Linder, et al. 2005), and MultiDivTime (Pirie, et al. 2006). As a result, in the past few years, NPRS has been largely abandoned in favor of its successor PL, and, more recently, UCLN. However, a certain number of studies found different trends, with NPRS estimating younger ages than PL (PLBP/Ref.), than clock-based LF estimates (Clement, et al. 2004), or than PL (PLBP/Ref.) and MultiDivTime (Bell and Donoghue (2005). It is worth noting that both studies, even though estimating younger ages with NPRS, try to accommodate their findings to the common expectation that NPRS should overestimate ages. Additionally, Paton et al. (2002) found divergence dates under PL (PLBP/Ref.) and NPRS to be very similar, Bremer et al. (2004) found small differences between PL (PLBP/Ref.) and NPRS estimates, while Su and Saunders (2009) found divergence times within the Annonaceae, estimated with uncorrelated lognormal (i.e., UCLD) relaxed molecular clock in BEAST, to differ slightly from previous age estimates based on NPRS and PL (PLBP/Ref.) methods. Our comparison of NPRS with other dating methods, in particular for nodes in the area of the tree without age constraints, reveals that NPRS estimated the oldest ages (as in dsA), but also ages that are younger than UCLN, older than PLBP and at the same time very similar to those of PLFB estimates (dsD).

PL, regardless of the cross-validation method applied, estimated younger ages than UCLN. In general, PLBP estimated the youngest ages (excluding PATHd8), while PLFB estimated ages similar to those of NPRS. Sanderson (2002) found that PL (PLBP /Ref.) always outperforms NPRS and subsequent studies comparing the two dating methods

became rare. PLFB (Near and Sanderson 2004) base the selection of the optimal model of molecular evolution on the fossil information. Although this allowed estimating optimal smoothing parameters through cross-validation for very large data sets, this method has rarely been applied in later studies. Today, PL is applied as often as Bayesian dating, still the cross-validation procedure stays computationally intensive, while the more commonly used branch-pruning cross-validation method is, additionally, not able to deal with large data sets. Anderson (2007) re-analysed the data set from Janssen and Bremer (2004) and obtained ages under PLFB analysis procedure that were highly similar to the ages found by NPRS in the original study. Similar to the trends recovered in our study, Magallón (2010) found PLFB to provide younger estimates than UCLN, while PLBP provided the youngest ages across the tree.

In all our data sets with comprehensive taxon and fossil sampling (dsA to dsD) PATHd8 consistently estimated the youngest ages. Assuming local molecular clock in particular tree segments, this dating method smoothens the mean substitution rates between sister taxa, consequently allowing for very fast analyses of large phylogenetic trees (Britton, et al. 2007). As data sets tend to increase in size, these attributes would make PATHd8 a primary candidate for dating analyses applying the DFC approach, but unfortunately there may be important concerns of bias for nodes with no adjacent fixed age nodes (Britton, et al. 2007, Svennblad 2008) and a recurrent underestimation of ages. Congruent with our observations, comparison of divergence times estimated by different dating methods in other studies reveals that PATHd8 estimates younger ages than PL (PLBP[Ref.]) (Ericson, et al. 2006, Brown, et al. 2008), NPRS (Britton, et al. 2007, Soltis, et al. 2008), PLFB and NPRS analyses (Anderson 2007, Britton, et al. 2007, Anderson and Janssen 2009), or in some cases largely correlated ages to those of PL (PLBP [Ref.]) and BEAST (Frajman, et al. 2009).

Dealing in our study with empirical data for which the true divergence times and the rates of molecular evolution are unknown, it is difficult to appraise between the methods that estimate different ages for the same nodes. Differences between the divergence time estimation methods used in this study are numerous and complex, related, among others to their functionality, assumptions and the priors they use (Pérez-Losada, et al. 2004). A general trend concerning the relative behavior and performance of the different analytical methods in a study setup corresponding to our DFC approach cannot be discerned, since more often deviations in the estimates seem to be based on the composition of data set in terms of taxon and fossil sampling rather than the analytical method applied.

Challenges with determining the optimal smoothing parameter for PL

At each increment, the difference between the smoothing parameter chosen by testing the entire range versus a limited range was much larger for PLBP than for PLFB. As a result, optimal smoothing parameters chosen through a PLBP cross-validation procedure,

while testing the whole range, estimated considerably older ages than when range was limited (Fig. 9). This pattern was observed for the area of the tree without fossil constraints, but also for the nodes in the fossil constrained area of the tree. The ages estimated with the smoothing parameters chosen through the PLFB cross-validations did not change considerably with the size of the range, or with the increments that were tested (Fig. 9).

Considering that the cross-validation procedure failed for a large number of data sets, we tested what the effect of applying a non-optimal smoothing parameter for a given data set by the analyses. This included applying a smoothing parameter that was estimated to be the optimal for the smaller sub-set on a larger data set; setting the smoothing parameter to an absolute, arbitrary value; or choosing a value consistent with range of values determined for the smaller data sets. A similar practice as been applied by a number of publications (e.g., Bremer, et al. 2004, Linder, et al. 2005). Ages estimated for dsA and dsD with the same smoothing parameter differed substantially (Fig. 10). Only few of the tested smoothing magnitudes, between λ_{10} = 2.17 and 2.24, estimated similar ages for all of the outlined nodes in both data sets (data shown for the CG node of Cactaceae, Fig. 10). Still, these smoothing parameters were much larger than the parameters chosen as optimal in both cross-validation types (i.e., λ_{10} = 0.43 for PLFB and λ_{10} = 1.75 for PLBP) or at any of the increment. Interestingly, all of the smoothing parameters smaller than λ_{10} = 2.17, when applied to dsA, estimated considerably older ages for the same node than when the same smoothing parameter was applied on dsD. In contrast, for smoothing parameters larger than λ_{10} = 2.24, age estimates for the same node of the dsA were considerably younger than estimates for dsD. This indicates that for a data set for which a cross-validation has failed, another smoothing parameter (either chosen approximately or estimated as being optimal for its sub-set) should never be applied, as there is a high risk of estimating largely discrepant divergence times. Indeed, this practice is flawed already by the definition of the cross-validation, as the chosen parameter will not be driven by the data itself (Sanderson 2002, Near and Sanderson 2004).

Choosing Appropriate Fossil Sampling in External Groups

Effect of Excluding the Maximum (Upper) Limit Constraint - A common problem in many studies seems to be associated with defining a realistic maximum age constraint (Benton and Donoghue 2007), and particularly the one that will not unrealistically overestimate ages of the deepest nodes. Smith et al. (2010) found that constraining the first appearance of tricolpate pollen as a fixed calibration to 125 Ma may underestimate the origin of eudicots, and consequently other age estimates relying on this constraint by some 3 to 22 Ma. Hug and Roger (2007) found that fixing the upper age of a tree leads to age estimates skewed towards ancient divergences. The only dating method that allows for age estimations without constraining the node of CG eudicots to an age of 125 Ma, and therefore defining the upper limit constraint for the whole data set, is UCLN, as all other methods we employed needed at

least one fixed calibration point. It resulted in overestimating ages in the three examined nodes from the fossil-constrained part of the tree (Fig. 4, Supplement S8), but only for the densely sampled data sets (i.e., dsI compared to dsA). Other examined nodes were estimated to very similar values in both data sets. These findings correspond to previous, and expected, observations that deeply nested nodes become older when a maximum age is not enforced (e.g. Gernandt, et al. 2008). This effect was, interestingly, not observed between the two data sets with the reduced taxon density (dsD and dsJ), where all of the examined nodes were estimated to very similar ages. In general, results yielded by our data sets with released upper age (dsI and dsJ) showed that our study essentially did not suffer from this effect, probably due to many calibration points that acted as bracketing boundaries.

Excluding Dubious Fossils from the Closest Fossil-Poor External group – We tested the effect of releasing three uncorroborated fossils that are placed closest to the cacti, as their proximity and (in)accuracy could have a significant impact on the reliability of estimated node ages (Yang and Yoder 2003, Yoder and Yang 2004). Interestingly, the observed effect in our study was however shown to be negligible for the densely sampled data set (dsK compared to dsA; Fig. 4, Supplement S8). Sampling taxa sparsely, in the data set with 29 fossils (dsL compared to dsD), underestimated ages in the area of the tree without fossil constraints, and the variance of the underestimation depended on the dating method used. The least sensitive method on eliminating the three uncorroborated fossils was PATHd8, regardless of the taxon sampling density, followed by UCLN, which underestimated ages in the sparsely sampled dsL. This underestimation was even more pronounced in NPRS, which in dsL estimated the youngest ages between all data sets with multiple fossil calibrations (i.e., dsA to dsL). It is difficult to distinguish if this was the result of the distance to the last calibrated node, or the inaccuracy of the release of the uncorroborated fossils, but it might suggest that in NPRS, when taxa are sampled sparsely, one of those two issues (or possibly both) are presumably correlated with the underestimating divergence times. The NPRS ages estimated in dsL were even younger from ages estimated in sparsely sampled data sets that apply the FGP approach (dsF) and the SC approach (dsH), suggesting that NPRS could be more sensitive to changes in the fossil assembly than to changes in the sampling approach. Both PLFB and PLBP estimated in dsL considerably younger ages for all nodes, and particularly for those that were further away from the fossil constrained area. Our findings, at least for NPRS, correspond to those of Linder et al. (2005), who recovered a significant, positive, linear relationship between the underestimation of a node and its distance from the calibration point in the sparsely sampled data sets, analyzed with NPRS and Bayesian methods. Still, Sauquet et al. (2012) found that removing one "risky" age constraint from the outgroup had a negligible effect on estimated ages (scenarios 3 and 6 in their study; which corresponds to our DFC approach)

Single vs Multiple Calibration Constraints – The long term debate surrounding the use of a single versus multiple fossil constraints in molecular dating analyses (e.g. Shaul and Graur

2002, Graur and Martin 2004, Hedges and Kumar 2004) is lately resolved in favor of the latter, as the use of a single fossil constraint generates highly unpredictable and variable estimates (Bremer, et al. 2004, Hug and Roger 2007). Yet, currently reported ages for many clades are still based on studies that used a single calibration constraint, and occasionally those ages are, nevertheless, used as the starting calibration for contemporary molecular dating studies (e.g. Franzke, et al. 2009). For all methods used in this study, regardless of the taxon sampling density or sampling approach (DFC, FGP, or SC), constraining a single calibration point at the root of the tree (dsM to dsR), compared to divergence times estimated for dsA and dsD, provided underestimated ages across all nodes in the tree (Fig. 5, Supplement S8), and once more, more complex methods (i.e., UCLN and PLBP, PLFB) proved to be more sensitive (Figs. 8a to 8h). The least sensitive method was PATHd8, the most sensitive was UCLN, while NPRS yielded slightly underestimated ages for the data sets with a dense taxon sampling dsM, dsO and dsQ). Sampling taxa sparsely (dsN, dsP, and dsR) in NPRS underestimated ages to various degrees. Curiously, almost identical ages were estimated for all nodes in the UCLN for dsM and dsQ – the former a data set with a single calibration point, constrained very distantly from the focal group (125 Ma at the CG eudicots), the latter a data set whose single constraint was phylogenetically much closer (100 Ma for CG Caryophyllales). We, however, believe it to be casual, as the SC approach (dsG) in general, underestimated ages. Using a single secondary point calibration (dsQ) likely underestimated ages even more. These ages fall only marginally within the range of the mean estimates of the data set dsM, and the conclusion that this suggest the insensitivity of the UCLN to the distance of the single constraint node is not valid. Sauquet et al. (2012) test the accuracy and the reliability of the secondary calibration approach by experimenting with a number of alternative secondary calibration settings. Settings in their scenarios 8a to 8e of their study are comparable with our single calibration constraint in dsQ and dsR. Similar to our findings, in each of their secondary calibration scenarios, Sauquet et al. (2012) obtained younger ages, concluding that it could lead to biased estimates.

Empirical Evaluation of Estimated Ages

A direct comparison of the divergence times estimated in this study is not straightforward: they have been produced under different methods and information provided by sampled densities of taxa and fossils, and often derived from trees with different topologies and branch lengths. However, as most dating studies aim at producing a "best estimate" for the age of a clade, it is arguably still valid (or biologically informative) to compare ages despite their being produced under different circumstances. We therefore report, compare and discuss the ages obtained for a set of clades identifiable in all of our investigated phylogenies: CG asterids, SL and CG Caryophyllales for the part of the tree constrained with fossils and SL and CG Portulacineae, SL and CG Cactaceae, SL and CG Cactoideae and SL and CG Trichocereinae for the part of the tree without age constraints

(Fig. 4). We do not consider the values of the UCLN estimate of dsB (because of mixing and convergence issues), nor of dsM to dsR (because we consider unjustifiable to use a single fossil calibration – the appearance of tricolpate pollen – as the only constraint in estimating the age of Cactaceae). We further do not discuss the values estimated by PATHd8 (which are however reported in Supplement S8 and in Fig. 4), because they are clearly biased for nodes with no adjacent fixed-age nodes (Svennblad 2008). Additionally, for the stem lineage and the crown group node of Caryophyllales we exclude values of the estimates from dsG and dsH, because the fixed age of CG Caryophyllales makes the node of SL Caryophyllales in these data sets to be the root node and concerns have been raised for estimating the age of root nodes (Sanderson 1997, Schenk and Hufford 2010).

Radically different age estimates have been proposed for the clades addressed in this study (Supplement S2), derived from different data sets, fossil calibrations, analytical methods and alternative topologies (Bromham, et al. 1999, Bremer, et al. 2004, Bell and Donoghue 2005, Bell, et al. 2005, Edwards and Hawkins 2007, Hug and Roger 2007). Our age estimates for Caryophyllales and its major lineages are only one example of the great extent of this variation (Fig. 4, Supplement S8), and we presume this to be a result of a combination of various factors influencing molecular dating (e.g., sampling approach, taxon sampling density, applied dating method and age constraints). For instance, published ages for the first appearance of Caryophyllales vary between 121 and 72 Ma (Supplement S2), which are in agreement with our estimates. According to previously reported ages, diversification of Caryophyllales started between 106 and 67 Ma, which is slightly to considerably younger than found in our analyses. The SL Cactaceae is dated in our study to an average of 52.5 (± 10.1) Ma, placing the first appearance of cacti to the early Eocene (i.e., Ypresian, 56.0 - 47.8 Ma; GSA Geological Time Scale, V. 4, 2012), however the oldest mean age of 71.1 Ma (UCLN, dsD) and the youngest of 32.5 Ma (UCLN, dsF), increasing the uncertainty in the placement of this event. The CG Cactaceae was dated to an average of 43.2 (± 12.7) Ma ago, assigning their radiation to the middle Eocene, still with the oldest mean estimate of 63.1 Ma (UCLN, dsD) and the youngest of 17.5 Ma (UCLN, dsF) providing a wider time span for this occurrence. Average estimations yielded by our study, especially for the origin of and the basal splits in Cactaceae are considerably older than those from the recent study of Arakaki et al. (2011), however do not question the contemporaneous and recent radiations suggested in that study, as most of the extant diversity of cacti recovered in our analyses occurred around the mid-Miocene and ages estimated within the Bayesian framework in our (UCLN, dsA) and that (MultiDivTime) study yielded comparable ages.

Conclusion and Perspectives

In this study we focused on estimating divergence times in a particularly challenging case when the taxonomic group of interest has a scarce or missing fossil record. We refer to it as the "molecular dating paradox" – taxonomic groups with an insufficient fossil record and, therefore, potentially benefitting most from the molecular dating are those that turn up to be most variable and inconsistent in estimated ages. In the center of our interest is the family Cactaceae - a prominent plant lineage that lacks a fossil record. We performed comparative studies investigating the effect of different sampling strategies, using different taxon sampling densities and analytical methods. We contribute new insights towards a deeper understanding for the wide range of divergent hypotheses on the origin of Cactaceae by classifying studies into one of the three sampling approaches we discern the DFC approach (i.e., distant fossil calibration), the FGP approach (i.e., focal group placeholder), and the SC approach (i.e., secondary calibration), or a combination thereof. Results yielded from our studies indicate that in cases when fossils are scarce, the DFC approach appears to be the most advanced approach, while the current trend towards using more sophisticated relaxed-clock methods such as UCLN and PL is very much dependent, and influenced, by the sampling approach and the taxon sampling density applied – contrary to simpler methods (i.e., NPRS or PATHd8), which are only rarely considered for age estimation studies.

A general trend concerning the relative behavior and performance of the different analytical methods in a study set-up corresponding to our DFC approach cannot be discerned, since more often deviations in the estimates seem to be related to the kind of data set used, rather than the analytical method applied. Further studies using uncorrelated relaxed-clock methods are needed to estimate the effect of taxon sampling within the framework of the three sampling approaches differentiated. Despite repeated criticism, the so-called "secondary calibration approach" is still widely used in estimating divergence times for focal groups with a scarce fossil record. As found in our study, this approach might yield fairly similar ages to those estimated in the DFC approach considering the similar "starting age" for the secondarily constrained node, favorably incorporating the full posterior distribution of the primary analysis; and few additional (i.e. "internal") fossil constraints.

We showed that the current molecular dating methods might still not be sophisticated enough to converge in results and to estimate consistently divergence times in groups of organisms with a scarce fossil record, especially if different sampling approaches and/or taxon sampling densities are applied. Indeed, as illustrated in Fig. 4 (see also Supplement S8), the full range of the CG age for Cactaceae estimated in our study extends from 63.1 and 17.5 Ma (mean ages in UCLN; dsD and dsF, respectively); 52.5 and 38.6 Ma (NPRS, dsG and dsL); 46.3 and 38.4 Ma (PLFB, dsD and dsL); 45.9 and 19.3 (PLBP, dsH and dsL); and when PATHd8 is considered 16.5 and 7.1 Ma (dsH and dsE). In addition, if 95% HPD for UCLN analyses are accounted for, age of divergence for the CG Cactaceae may reach as much as 83.7 Ma (dsD). This makes clear that studies using a single molecular dating

method on the basis of a single data set do not investigate the full range of variance of age estimations that can be produced: their confidence intervals of node ages may produce a false sense of certainty, as they only incorporate a small proportion of the full underlying uncertainty. We suggest to comparatively investigate several different sets of input data, differing for instance in the sampling approach and the taxon sampling density, in particular when more sophisticated analytical methods, such as UCLN are applied

In summary, we show that there is a plethora of factors affecting the results in molecular dating analyses. Only by performing alternative analyses, changing one variable at a time, is it possible to assess their quantitative influence on node ages, but even then there is an uneven distribution of errors across the tree (depending e.g. on the distance to constrained nodes and topological and branch length uncertainty). A comfortable option for researchers interested in the evolution of fossil-deprived taxa would be not to estimate divergence times at all. We believe, however, that even given all the uncertainty associated with dating such taxa, there is a clear benefit in producing well-developed temporal hypotheses. We have outlined many dangers on the path to successful divergence time estimates and where possible, provided general recommendations helping to improve the latter. Nevertheless, the bottom line is that researchers need to carefully assess the robustness of their results given their data and methodology, and more importantly, commit themselves to revise their theories in light of further evidence.

Supplementary Documents

Supplementary documents are attached in this manuscript at the end of the present chapter.

Acknowledgments

We would like to thank to Aelys Humphreys and Michael D. Pirie for helpful discussions, and Jurriaan M. de Vos and Richard Carter for constructive comments on an earlier draft of this paper. Merran Matthews is thanked for being a perpetuator and an endless source of motivation through this project.

A.A. is supported by grants from the Swedish Research Council and the European Research Council (ERC StG #311024).

References

Anderson CL. 2007. Dating Divergence Times in Phylogenies. Faculty of Science and Technology 322. Uppsala, Uppsala Universitet, p. 79.

Anderson CL, Bremer K, Friis EM. 2005. Dating phylogenetically basal eudicots using *rbcL* sequences and multiple fossil reference points. American Journal of Botany, 92:1737-1748.

Anderson CL, Janssen T. 2009. Monocots. The TimeTree of Life, Oxford University Press, New York, p. 551.

Applequist WL, Wallace RS. 2001. Phylogeny of the portulacaceous cohort based on ndhF sequence data. Syst. Bot., 26:406-419.

Arakaki M, Christin PA, Nyffeler R, Lendel A, Eggli U, Ogburn RM, Spriggs E, Moore MJ, Edwards EJ. 2011. Contemporaneous and recent radiations of the world's major succulent plant lineages. Proc. Natl. Acad. Sci. U. S. A., 108:8379-8384.

Baldwin BG, Sanderson MJ. 1998. Age and rate of diversification of the Hawaiian silversword alliance (Compositae). Proc. Natl. Acad. Sci. U. S. A., 95:9402-9406.

Barthlott W. 1983. Biogeography and evolution in neo- and palaeotropical Rhipsalinae. In K. Kubitzki [ed.], Dispersal and distribution. In: Kubitzki IK editor. Dispersal and distribution. Hamburg, Germany., Verlag Paul Parey, p. 241-248.

Barthlott W, Hunt DR. 1993. The families and genera of vascular plants. In: Kubitzki, K & al. (eds.). Berlin (D) etc., Springer-Verlag, p. 161-197, ills.

Becker HF. 1962. Reassignment of Eopuntia to Cyperacites. Bull. Torrey Bot. Club, 89:319-330.

Bell CD, Donoghue MJ. 2005. Dating the Dipsacales: Comparing models, genes, and evolutionary implications. American Journal of Botany, 92:284-296.

Bell CD, Soltis DE, Soltis PS. 2005. The age of the angiosperms: A molecular timescale without a clock. Evolution, 59:1245-1258.

Bell CD, Soltis DE, Soltis PS. 2010. The age and diversification of the angiosperms re-revisited. American Journal of Botany, 97:1296-1303.

Benton MJ. 2000. Vertebrate Palaeontology. Blackwell Science.

Benton MJ, Donoghue PCJ. 2007. Paleontological evidence to date the tree of life (vol 24, pg 26, 2007). Mol. Biol. Evol., 24:889-891.

Bremer B, Bremer K, Heidari N, Erixon P, Olmstead RG, Anderberg AA, Kallersjo M, Barkhordarian E. 2002. Phylogenetics of asterids based on 3 coding and 3 non-coding chloroplast DNA markers and the utility of non-coding DNA at higher taxonomic levels. Mol. Phylogenet. Evol., 24:274-301.

Bremer K, Friis EM, Bremer B. 2004. Molecular phylogenetic dating of asterid flowering plants shows early Cretaceous diversification. Syst. Biol., 53:496-505.

Britton T, Anderson CL, Jacquet D, Lundqvist S, Bremer K. 2007. Estimating divergence times in large phylogenetic trees. Syst. Biol., 56:741-752.

Bromham L, Phillips MJ, Penny D. 1999. Growing up with dinosaurs: molecular dates and the mammalian radiation. Trends Ecol. Evol., 14:113-118.

Brown JW, Rest JS, Garcia-Moreno J, Sorenson MD, Mindell DP. 2008. Strong mitochondrial DNA support for a Cretaceous origin of modern avian lineages. BMC Biol., 6.

Buchheim G. 1964. Cactales. In: Melchior H editor. A. Engler's Syllabus der Pflanzenfamilien. Berlin, Gebrüder Borntraeger, ed. 12, 2:, p. 102-108.

Buschiazzo E, Ritland C, Bohlmann J, Ritland K. 2012. Slow but not low: genomic comparisons reveal slower evolutionary rate and higher dN/dS in conifers compared to angiosperms. BMC Evol. Biol., 12.

Butterworth C, Edwards EJ. 2008. Investigating *Pereskia* and the Earliest Divergences In Cactaceae. Haseltonia:46-53.

Chamberland M. 1997. Systematics of the Echinocactus polycephalus complex (Cactaceae). Syst. Bot., 22:303-313.
Chaney RW. 1944. A fossil cactus from the Eocene of Utah. American Journal of Botany, 31:507-528.
Clement WL, Tebbitt MC, Forrest LL, Blair JE, Brouillet L, Eriksson T, Swensen SM. 2004. Phylogenetic position and biogeography of Hillebrandia sandwicensis (Begoniaceae): A rare Hawaiian relict. American Journal of Botany, 91:905-917.
Cody S, Richardson JE, Rull V, Ellis C, Pennington RT. 2010. The Great American Biotic Interchange revisited. Ecography, 33:326-332.
Conti E, Eriksson T, Schonenberger J, Sytsma KJ, Baum DA. 2002. Early tertiary out-of-India dispersal of Crypteroniaceae: Evidence from phylogeny and molecular dating. Evolution, 56:1931-1942.
Cuénoud P, Savolainen V, Chatrou LW, Powell M, Grayer RJ, Chase MW. 2002. Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid rbcL, atpB, and matK DNA sequences. American Journal of Botany, 89:132-144.
Doyle JA. 2012. Molecular and Fossil Evidence on the Origin of Angiosperms. Annu. Rev. Earth Planet. Sci., 40:301-326.
Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. PLoS. Biol., 4:699-710.
Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol., 7.
Edwards D, Hawkins JA. 2007. Are Cape floral clades the same age? Contemporaneous origins of two lineages in the genistoids s.l. (Fabaceae). Mol. Phylogenet. Evol., 45:952-970.
Edwards EJ, Nyffeler R, Donoghue MJ. 2005. Basal cactus phylogeny: implications of *Pereskia* (Cactaceae) paraphyly for the transition to the cactus life form. American Journal of Botany, 92:1177-1188.
Ericson PGP, Anderson CL, Britton T, Elzanowski A, Johansson US, Kallersjo M, Ohlson JI, Parsons TJ, Zuccon D, Mayr G. 2006. Diversification of Neoaves: integration of molecular sequence data and fossils. Biol. Lett., 2:543-U541.
Felsenstein J. 1981. Evolutionary Trees from DNA-Sequences - a Maximum-Likelihood Approach. J. Mol. Evol., 17:368-376.
Felsenstein J. 1985. Confidence-Limits on Phylogenies - an Approach Using the Bootstrap. Evolution, 39:783-791.
Fior S, Li M, Oxelman B, Viola R, Hodges SA, Ometto L, Varotto C. 2013. Spatiotemporal reconstruction of the Aquilegia rapid radiation through next-generation sequencing of rapidly evolving cpDNA regions. . New Phytol., 198:579-592.
Frajman B, Eggens F, Oxelman B. 2009. Hybrid Origins and Homoploid Reticulate Evolution within Heliosperma (Sileneae, Caryophyllaceae)-A Multigene Phylogenetic Approach with Relative Dating. Syst. Biol., 58:328-345.
Franzke A, German D, Al-Shehbaz IA, Mummenhoff K. 2009. Arabidopsis family ties: molecular phylogeny and age estimates in Brassicaceae. Taxon, 58:425-437.
Gentry AH. 1982. Neotropical Floristic Diversity: Phytogeographical Connections Between Central and South America, Pleistocene Climatic Fluctuations, or an Accident of the Andean Orogeny? Annals of the Missouri Botanical Garden, 69:557-593.
Gernandt DS, Magallon S, Lopez GG, Flores OZ, Willyard A, Liston A. 2008. Use of Simultaneous Analyses to Guide Fossil-Based Calibrations of Pinaceae Phylogeny. Int. J. Plant Sci., 169:1086-1099.
Gibson AC, Nobel PS. 1986. The cactus primer. Cambridge (US:MA)/London (GB), Harvard University Press.
Gillespie JH. 1991. The causes of molecular evolution., Oxford: Oxford University Press.
Goldblatt P, Savolainen V, Porteous O, Sostaric I, Powell M, Reeves G, Manning JC, Barraclough TG, Chase MW. 2002. Radiation in the Cape flora and the phylogeny of

peacock irises Moraea (Iridaceae) based on four plastid DNA regions. Mol. Phylogenet. Evol., 25:341-360.
Goodall-Copestake WP, Harris DJ, Hollingsworth PM. 2009. The origin of a mega-diverse genus: dating Begonia (Begoniaceae) using alternative datasets, calibrations and relaxed clock methods. Bot. J. Linnean Soc., 159:363-380.
Graham A, Gregory-Wodzicki KM, Wright KL. 2001. Studies in neotropical paleobotany. XV. A Mio-Pliocene palynoflora from the Eastern Cordillera, Bolivia: Implications for the uplift history of the Central Andes. American Journal of Botany, 88:1545-1557.
Graur D, Martin W. 2004. Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. Trends Genet., 20:80-86.
Gustafsson ALS, Verola CF, Antonelli A. 2010. Reassessing the temporal evolution of orchids with new fossils and a Bayesian relaxed clock, with implications for the diversification of the rare South American genus Hoffmannseggella (Orchidaceae: Epidendroideae). BMC Evol. Biol., 10.
Heads M. 2005. Dating nodes on molecular phylogenies: a critique of molecular biogeography. Cladistics, 21:62-78.
Heads M. 2011. Old taxa on young islands: a critique of the use of island age to date island-endemic clades and calibrate phylogenies. Syst Biol, 60:204-218.
Heath TA, Hedtke SM, Hillis DM. 2008a. Taxon sampling and the accuracy of phylogenetic analyses. J. Syst. Evol., 46:239-257.
Heath TA, Zwickl DJ, Kim J, Hillis DM. 2008b. Taxon Sampling Affects Inferences of Macroevolutionary Processes from Phylogenetic Trees. Syst. Biol., 57:160-166.
Hedges SB, Kumar S. 2004. Precision of molecular time estimates. Trends Genet., 20:242-247.
Hershkovitz MA, Zimmer EA. 1997. On the evolutionary origins of the cacti. Taxon, 46:217-231.
Ho SYW. 2007. Calibrating molecular estimates of substitution rates and divergence times in birds. J. Avian Biol., 38:409-414.
Ho SYW, Phillips MJ. 2009. Accounting for Calibration Uncertainty in Phylogenetic Estimation of Evolutionary Divergence Times. Syst. Biol., 58:367-380.
Ho SYW, Saarma U, Barnett R, Haile J, Shapiro B. 2008. The Effect of Inappropriate Calibration: Three Case Studies in Molecular Ecology. PLoS One, 3.
Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics, 17:754-755.
Hug LA, Roger AJ. 2007. The impact of fossils and taxon sampling on ancient molecular dating analyses. Mol. Biol. Evol., 24:1889-1897.
Hugall AF, Foster R, Lee MSY. 2007. Calibration choice, rate smoothing, and the pattern of tetrapod diversification according to the long nuclear gene RAG-1. Syst. Biol., 56:543-563.
Hunt DR, Taylor NP. 1990. The genera of the Cactaceae: progress towards consensus. Bradleya, 8:85-107.
Janssen T, Bremer K. 2004. The age of major monocot groups inferred from 800+ rbcL sequences. Bot. J. Linnean Soc., 146:385-398.
Janssens SB, Knox EB, Huysmans S, Smets EF, Merckx V. 2009. Rapid radiation of Impatiens (Balsaminaceae) during Pliocene and Pleistocene: Result of a global climate change. Mol. Phylogenet. Evol., 52:806-824.
Jian SG, Soltis PS, Gitzendanner MA, Moore MJ, Li R, Hendry TA, Qiu YL, Dhingra A, Bell CD, Soltis DE. 2008. Resolving an ancient, rapid radiation in Saxifragales. Syst. Biol., 57:38-57.
Kumar S, Filipski A, Swarna V, Walker A, Hedges SB. 2005. Placing confidence limits on the molecular age of the human-chimpanzee divergence. Proc. Natl. Acad. Sci. U. S. A., 102:18842-18847.

Landrum JV. 2002. Four succulent families and 40 million years of evolution and adaptation to xeric environments: What can stem and leaf anatomical characters tell us about their phylogeny? Taxon, 51:463-473.
Lee MSY. 1999. Molecular clock calibrations and metazoan divergence dates. J. Mol. Evol., 49:385-391.
Linder HP, Hardy CR, Rutschmann F. 2005. Taxon sampling effects in molecular clock dating: An example from the African Restionaceae. Mol. Phylogenet. Evol., 35:569-582.
Maddison DR, Maddison WP. 2000. MacClade. Sunderland, MA, Sinauer Associates Inc.
Magallón S. 2004. Dating lineages: molecular and paleontological approaches to the temporal framework of clades. Int. J. Plants Sci., 156.
Magallón S. 2010. Using Fossils to Break Long Branches in Molecular Dating: A Comparison of Relaxed Clocks Applied to the Origin of Angiosperms. Syst. Biol., 59:384-399.
Magallón S, Castillo A. 2009. Angiosperm Diversification through Time. American Journal of Botany, 96:349-365.
Mauseth JD. 1990. Continental drift, climate and the evolution of the cacti. Cactus and succulent journal (U.S.A.), 62:302-308.
Michalak I, Zhang LB, Renner SS. 2010. Trans-Atlantic, trans-Pacific and trans-Indian Ocean dispersal in the small Gondwanan Laurales family Hernandiaceae. J. Biogeogr., 37:1214-1226.
Milne RI. 2009. Effects of taxon sampling on molecular dating for within-genus divergence events, when deep fossils are used for calibration. J. Syst. Evol., 47:383-401.
Moore MJ, Soltis PS, Bell CD, Burleigh JG, Soltis DE. 2010. Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. Proc. Natl. Acad. Sci. U. S. A., 107:4623-4628.
Müller K. 2005. SeqState -primer design and sequence statistics for phylogenetic DNA data sets. Applied Bioinformatics, 4:65-69.
Near TJ, Benard MF. 2004. Rapid allopatric speciation in logperch darters (Percidae: Percina). Evolution, 58:2798-2808.
Near TJ, Bossu CM, Bradburd GS, Carlson RL, Harrington RC, Hollingsworth PR, Jr., Keck BP, Etnier DA. 2011. Phylogeny and temporal diversification of darters (Percidae: Etheostomatinae). Syst Biol, 60:565-595.
Near TJ, Sanderson MJ. 2004. Assessing the quality of molecular divergence time estimates by fossil calibrations and fossil-based model selection. Philos. Trans. R. Soc. Lond. Ser. B-Biol. Sci., 359:1477-1483.
Nie ZL, Wen J, Azurna H, Qiu YL, Sun H, Meng Y, Sun WB, Zimmer EA. 2008. Phylogenetic and biogeographic complexity of Magnoliaceae in the Northern Hemisphere inferred from three nuclear data sets. Mol. Phylogenet. Evol., 48:1027-1040.
Nyffeler R. 2002. Phylogenetic relationships in the cactus family (Cactaceae) based on evidence from trnK/matK and trnL-trnF sequences. American Journal of Botany, 89:312-326.
Nyffeler R. 2007. The closest relatives of cacti: Insights from phylogenetic analyses of chloroplast and mitochondrial sequences with special emphasis on relationships in the tribe Anacampseroteae. American Journal of Botany, 94:89-101.
Nyffeler R, Eggli U. 2010a. Disintegrating Portulacaceae: A new familial classification of the suborder Portulacineae (Caryophyllales) based on molecular and morphological data. Taxon, 59:227-240.
Nyffeler R, Eggli U. 2010b. A farewell to dated ideas and concepts: molecular phylogenetics and a revised suprageneric classification of the family Cactaceae. Schumannia, 6:109-149.
Nylinder S, Swenson U, Persson C, Janssens SB, Oxelman B. 2012. A dated species-tree approach to the trans-Pacific disjunction of the genus Jovellana (Calceolariaceae, Lamiales). Taxon, 61:381-391.

Ocampo G, Columbus JT. 2010. Molecular phylogenetics of suborder Cactineae (Caryophyllales), including insights into photosynthetic diversification and historical biogeography. American Journal of Botany, 97:1827-1847.
Paton T, Haddrath O, Baker AJ. 2002. Complete mitochondrial DNA genome sequences show that modern birds are not descended from transitional shorebirds. Proceedings of the Royal Society of London. Series B: Biological Sciences, 269:839-846.
Pérez-Losada M, Høeg JT, Crandall KA. 2004. Unraveling the Evolutionary Radiation of the Thoracican Barnacles Using Molecular and Morphological Evidence: A Comparison of Several Divergence Time Estimation Approaches. Syst. Biol., 53:244-264.
Pfeil BE, Crisp MD. 2008. The Age and Biogeography of Citrus and the Orange Subfamily (Rutaceae: Aurantioideae) in Australasia and New Caledonia. American Journal of Botany, 95:1621-1631.
Pirie MD, Chatrou LW, Mols JB, Erkens RHJ, Oosterhof J. 2006. 'Andean-centred' genera in the short-branch clade of Annonaceae: testing biogeographical hypotheses using phylogeny reconstruction and molecular dating. J. Biogeogr., 33:31-46.
Pitts JP, Wilson JS, von Dohlen CD. 2010. Evolution of the nocturnal Nearctic Sphaeropthalminae velvet ants (Hymenoptera: Mutillidae) driven by Neogene orogeny and Pleistocene glaciation. Mol. Phylogenet. Evol., 56:134-145.
Posada D, Crandall KA. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics, 14:817-818.
Prance GT. 1994. A Comparison of the Efficacy of Higher Taxa and Species Numbers in the Assessment of Biodiversity in the Neotropics. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 345:89-99.
Pulquerio MJF, Nichols RA. 2007. Dates from the molecular clock: how wrong can we be? Trends Ecol. Evol., 22:180-184.
Reisz RR, Muller J. 2004. Molecular timescales and the fossil record: a paleontological perspective. Trends Genet., 20:237-241.
Robinson M, Gouy M, Gautier C, Mouchiroud D. 1998. Sensitivity of the relative-rate test to taxonomic sampling. Mol. Biol. Evol., 15:1091-1098.
Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics, 19:1572-1574.
Ronquist F, Klopfstein S, Vilhelmsen L, Schulmeister S, Murray DL, Rasnitsyn AP. 2012a. A Total-Evidence Approach to Dating with Fossils, Applied to the Early Radiation of the Hymenoptera. Syst. Biol.
Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012b. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. Syst. Biol., 61:539-542.
Rutschmann F. 2006. Molecular dating of phylogenetic trees: A brief review of current methods that estimate divergence times. Divers. Distrib., 12:35-48.
Rutschmann F, Eriksson T, Abu Salim K, Conti E. 2007. Assessing calibration uncertainty in molecular dating: The assignment of fossils to alternative calibration points. Syst. Biol., 56:591-608.
Sanderson MJ. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. Mol. Biol. Evol., 14:1218-1231.
Sanderson MJ. 2002. Estimating absolute rates of molecular evolution and divergence times: A penalized likelihood approach. Mol. Biol. Evol., 19:101-109.
Sanderson MJ. 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics, 19:301-302.
Sanderson MJ. 2006. r8s version 1.71. Analysis of rates ("r8s") of evolution., Section of Evolution and Ecology, University of California, Davis.
Sauquet H, Ho SYW, Gandolfo MA, Jordan GJ, Wilf P, Cantrill DJ, Bayly MJ, Bromham L, Brown GK, Carpenter RJ*, et al.* 2012. Testing the impact of calibration on molecular

divergence times using a fossil-rich group: the case of nothofagus (fagales). Syst. Biol., 61:289-313.
Schenk JJ, Hufford L. 2010. Effects of Substitution Models on Divergence Time Estimates: Simulations and an Empirical Study of Model Uncertainty Using Cornales. Syst. Bot., 35:578-592.
Shaul S, Graur D. 2002. Playing chicken (*Gallus gallus*): methodological inconsistencies of molecular divergence date estimates due to secondary calibration points. Gene, 300:59-61.
Simmons MP, Ochoterena H. 2000. Gaps as characters in sequence-based phylogenetic analyses. Syst. Biol., 49:369-381.
Smith AB. 1994. Systematics and the Fossil Record. Blackwell Scientific Publications.
Smith SA, Beaulieu JM, Donoghue MJ. 2010. An uncorrelated relaxed-clock analysis suggests an earlier origin for flowering plants. Proc. Natl. Acad. Sci. U. S. A., 107:5897-5902.
Soares AER, Schrago CG. 2012. The influence of taxon sampling and tree shape on molecular dating: an empirical example from Mammalian mitochondrial genomes. Bioinformatics and biology insights, 6:129-143.
Soltis DE, Bell CD, Kim S, Soltis PS. 2008. Origin and Early Evolution of Angiosperms. Year in Evolutionary Biology 2008. Oxford, Blackwell publishing, p. 3-25.
Soltis DE, Gitzendanner MA, Soltis PS. 2007. A 567-taxon data set for angiosperms: The challenges posed by bayesian analyses of large data sets. Int. J. Plant Sci., 168:137-157.
Su YCF, Saunders RMK. 2009. Evolutionary divergence times in the Annonaceae: evidence of a late Miocene origin of Pseuduvaria in Sundaland with subsequent diversification in New Guinea. BMC Evol. Biol., 9.
Svennblad B. 2008. Consistent Estimation of Divergence Times in Phylogenetic Trees with Local Molecular Clocks. Syst. Biol., 57:947-954.
Taylor JW, Berbee ML. 2006. Dating divergences in the Fungal Tree of Life: review and new analyses. Mycologia, 98:838-849.
The Angiosperm Phylogeny G. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. Bot. J. Linnean Soc., 161:105-121.
Thorne JL, Kishino H. 2002. Divergence time and evolutionary rate estimation with multilocus data. Syst. Biol., 51:689-702.
van Tuinen M, Hadly EA. 2004. Calibration and error in placental molecular clocks: A conservative approach using the cetartiodactyl fossil record. J. Hered., 95:200-208.
van Tuinen M, Hedges SB. 2001. Calibration of avian molecular clocks. Mol. Biol. Evol., 18:206-213.
van Tuinen M, Hedges SB. 2004. The effect of external and internal fossil calibrations on the avian evolutionary timescale. J. Paleontol., 78:45-50.
Welch JJ, Bromham L. 2005. Molecular dating when rates vary. Trends Ecol. Evol., 20:320-327.
Wikstrom N, Savolainen V, Chase MW. 2001. Evolution of the angiosperms: calibrating the family tree. Proc. R. Soc. Lond., 268:2211-2220.
Wilkinson RD, Steiper ME, Soligo C, Martin RD, Yang ZH, Tavare S. 2011. Dating Primate Divergences through an Integrated Analysis of Palaeontological and Molecular Data. Syst. Biol., 60:16-31.
Xiang QY, Thomas DT, Xiang QP. 2011. Resolving and dating the phylogeny of Cornales - Effects of taxon sampling, data partitions, and fossil calibrations. Mol. Phylogenet. Evol., 59:123-138.
Yang ZH, Yoder AD. 2003. Comparison of likelihood and Bayesian methods for estimating divergence times using multiple gene loci and calibration points, with application to a radiation of cute-looking mouse lemur species. Syst. Biol., 52:705-716.

Yoder AD, Yang ZH. 2004. Divergence dates for Malagasy lemurs estimated from multiple gene loci: geological and evolutionary context. Mol. Ecol., 13:757-773.
Zhang M-L, Fritsch PW. 2010. Evolutionary response of *Caragana* (Fabaceae) to Qinghai-Tibetan Plateau uplift and Asian interior aridification. Plant Syst. Evol., 288:191-199.

Figures and Tables

Figure 1. Three different strategies employed in estimating divergence times for a focal group with a scarce fossil record. Light and dark grey shadings mark the columns corresponding to Caryophyllales and Cactaceae, respectively (similar to shadings in the Figs. 2 and 3).

Figure 2. Summary phylogram of the selected tree topology derived from the Bayesian phylogenetic analyses of dsA used for age estimation. The tree contains 460 terminals and depicts relationships among major clades of eudicots. Outgroup taxa (Ceratophyllales) are not shown. The area not shaded on the tree corresponds to the eudicots, excluding Caryophyllales. Light gray shading marks the taxa sampled in Caryophyllales, while dark grey shading marks taxa belonging to Cactaceae. The line across the topology delimits position of the fossil constraints that are placed topologically closest to the cacti. Major clades are outlined above the tree, while many of the plant families are named next to their corresponding placement. Where possible, clades are collapsed to the level of the genus or, within Cactaceae, to the lower taxonomic rank. The size of the triangle is proportional to the number of taxa sampled within the clade. Numbers in circles indicate fossil constraints as listed in the Supplement S1. Fossils assigned to the stem lineage are shown arbitrarily along the branch separating the most recent common ancestor and node subtending all descendant taxa, while fossils assigned to the crown group node are shown on the node subtending all descendant taxa. Arrows indicate two main nodes of our interest. Node marked with a star represents the node at which the age was fixed with the secondary calibration constraint in data sets dsG, dsH, dsQ and dsR.

Figure 3. Dated phylogenetic tree of the full data set (dsA) produced through UCLN analyses, showing relationships among major clades of the eudicots and their divergence ages. The line across the topology delimits position of the fossil constraint that is placed topologically closest to the cacti. Error bars around nodes correspond to 95% HPD of divergence times and are shown only for a selected number of nodes. Other conventions as in Fig. 2. Age estimates for nodes marked are further compared and discussed (see Figs. 4 to 8).

Figure 4. Divergence age estimates obtained for different data sets, with different dating methods. Estimated ages of eleven selected nodes, obtained for different data sets and with different dating methods. Each "point" estimate of age represents the

mean (NPRS, PLFB, PLBP, PATHd8) or a mean with the associated error of 95% HPD (UCLN). Light and dark grey shadings mark the nodes belonging to Caryophyllales and Cactaceae, respectively, as in Figs. 2 and 3. The line across the topology indicates the position of the fossil constraint that is topologically closest to the cacti (i.e., fossil number 32 in the Supplement S1), placed on the figure according to its temporal position along the x-axis and its approximate topological placement along the y-axis. Average age estimates represent an average value of all estimated ages for a certain node, including all of the data sets (except for the dsM to dsR, with a single calibration point) and all dating methods investigated (except for PATHd8 and BEAST analysis for dsB). Additionally, in calculating the average age for the stem lineage and the crown group node of Caryophyllales we excluded values of the estimates from the data sets dsG and dsH. UCLN = uncorrelated lognormal with the 95% HPD; NPRS = nonparametric rate smoothing; PLFB = penalized likelihood with fossil-based rate smoothing; PLBP = penalized likelihood with branch-pruning rate smoothing; * = secondary calibration point; Avr. age = average age; Ma = million years ago.

Figure 5. Comparison of divergence times for eleven nodes estimated from three different sampling approaches (DFC, FGP, and SC), according to the dating method and the taxon sampling density applied. Each "point" estimate of age represents the mean (NPRS, PLFB, PLBP, PATHd8) or a mean with the associated error of 95% HPD (UCLN). Light and dark grey shadings embed the taxa belonging to the Caryophyllales and Cactaceae, respectively, associating to the similar shadings in the Figs 2 and 3. The fossil constraint that is topologically closest to cacti (fossil number 32 in the Supplement S1) is shown placed on the graph according to its temporal position along the x-axis and its approximate topological placement along the y-axis. Ma = million years ago; 2˚CalP = secondary calibration point.

Figure 6. Comparison of divergence times for eleven nodes estimated from data sets with different sampling density, according to the sampling approaches (i.e., DFC, FGP, and SC) and the dating methods applied. Each "point" estimate of age represents the mean (NPRS, PLFB, PLBP, PATHd8) or a mean with the associated error of 95% HPD (UCLN). Light and dark grey shadings embed the taxa belonging to the Caryophyllales and Cactaceae, respectively, associating to the similar shadings in the Figs 2 and 3. Fossil constraint that is topologically closest to the cacti (fossil number 32 in the Supplement S1) is shown placed on the graph according to its temporal position along the x axis and its approximate topological placement along the y axis. Curve on the right part on the graph outlines the change in the

coefficient of variation along the nodes. Ma = million years ago; 2˚CalP = secondary calibration point.

Figure 7. Comparison of divergence times for eleven nodes estimated with different dating methods, according to the sampling density applied. Each "point" estimate of age represents the mean (NPRS, PLFB, PLBP, PATHd8) or a mean with the associated error of 95% HPD (UCLN). Light and dark grey shadings embed the taxa belonging to the Caryophyllales and Cactaceae, respectively, associating to the similar shadings in the Figs 2 and 3. Fossil constraint that is topologically closest to the cacti (fossil number 32 in the Supplement S1) is shown placed on the graph according to its temporal position along the x axis and its approximate topological placement along the y axis. Curve on the right part on the graph outlines the change in the coefficient of variation along the nodes. Ma = million years ago; 2˚CalP = secondary calibration point.

Figure 8. Effect of estimating divergence times with a single calibration constraint. Comparison of divergence time estimates for selected nodes obtained by analysing data sets dsM to dsR with different molecular dating methods and sampling density, compared to data sets dsA and dsD. Each "point" estimate of age represents the mean (NPRS, PLFB, PLBP, PATHd8) or a mean with the associated error of 95% HPD (UCLN). Light and dark grey shadings embed the taxa belonging to the Caryophyllales and Cactaceae, respectively, associating to the similar shadings in the Figs 2 and 3. Fossil constraint that is topologically closest to the cacti, fossil number 32 (for the data sets dsA and dsD), fossil number 1 (for the data sets dsM to dsP) and the secondary calibration point (2˚CalP) for the data sets dsQ and dsR are shown, placed on the figure according to their temporal position along the x axis and their approximate topological placement along the y axis. Numbering of fossils corresponds to the one in the Supplement S1. UCLN = uncorrelated lognormal; NPRS = nonparametric rate smoothing; PLFB = penalized likelihood with fossil-based rate smoothing; PLBP = penalized likelihood with branch-pruning rate smoothing; Ma = million years ago.

Figure 9. Comparison of divergence times estimated for CG Cactaceae in penalized likelihood, using optimal smoothing parameters that were chosen through two different smoothing optimization criteria, testing the entire range versus limited ranges added in consecutive steps. Ma = million years ago.

Figure 10. Comparison of divergence time estimated for CG Cactaceae in penalized likelihood, resulting from applying same rate smoothing parameters on data sets dsA and dsD. Ma = million years ago.

Table 1. Description of data sets used in the dating analyses. A list of data sets and short description is outlined. Light and dark grey shadings mark the columns corresponding to the Caryophyllales and Cactaceae, respectively (similar shading adopted in Figs. 2 and 3).

Online Supplement

Supplement S1. Fossil constraints. Fossils used as age constraints are listed in the order that corresponds to the numbering of fossil constraints in Fig. 2. Two-letter code prior to the name of the node refers to how the fossil was assigned: along the stem lineage (SL) or at the crown group (CG). * SL of higher Caryophyllaceae includes the tribes Alsineae, Arenarieae, Eremogoneae, Sileneae, Caryophylleae according to Harbaugh et al. 2010 Int. J. Pl. Sci. 171: 185-198, 2010.

Supplement S2. Ages of selected nodes reported in the literature. Studies referring to the age of Cactaceae and Caryophyllales are listed and their ages are noted in millions of years. SL = stem lineage; CG = crown group; UCLN = an uncorrelated lognormal relaxed-clock model, with confidence interval values in brackets; NPRS = nonparametric rate smoothing method; PLFB = penalized likelihood with fossils-based rate smoothing; PLBP = penalized likelihood with branch-pruning rate smoothing; PL= penalized likelihood (smoothing method is not specified); Average age = average value excluding PATHd8, UCLN values for dsB, and data sets I and J; SD = standard deviation; [ ] = values of the six analyses; CI = confidence interval; UCLN# = values from multiple-fossil analyses that treated all fossils as exponential or lognormal priors are shown together; NPRS*; PATHd8* = range between the lowest and the highest estimate of analyses with a single fixed and multiple calibration analyses is shown; other" = comparing ITS homogeneity with different lineages of angiosperms.

Supplement S3. List of taxa included in this study and their GenBank accession numbers.

Supplement S4. Taxon sampling in the dsB, dsC and dsD. Phylogram derived from dsA with preserved taxa in black and pruned taxa in red. S4a illustrates phylogram consistent with data set dsB, S4b illustrates phylogram consistent with data set dsC, and S4c illustrates phylogram consistent with data set dsD.

Supplement S5. Major clades, number of taxa in the major clades according to the data sets and references for their phylogenetic placements. Major clades are listed according to their appearance in Fig. 2, reading from the top down. Names of data sets correspond to those in the Table 1. Lines of the table without coloration outline clades that belong to Distant Fossil-Rich external group (i.e. eudicots excluding Caryophyllales). The light and dark grey shadings embed the clades belonging to Caryophyllales and Cactaceae, respectively, associating to the similar shadings in the Figs 2 and 3. The last two columns list the references used to infer the phylogenetic placement of clades and the relationships within clades.

Supplement S6. Bayesian 50 % majority rule consensus phylogram.

Supplement S7. MP strict consensus tree.

Supplement S8. Diversification time estimates of selected clades. Stem lineage and crown node ages in million years (Ma) are outlined for 11 selected nodes obtained by estimating ages in 18 data sets with 4 relaxed clocks that implement among-lineage rate heterogeneity differently: Bayesian uncorrelated method (UCLN), nonparametric rate smoothing (NPRS), penalized likelihood (using 2 different rate smoothing optimization criteria) and PATHd8. Values represent the mean (NPRS, PLFB, PLBP, PATHd8) or a mean with the associated error of 95% HPD (UCLN). The light and dark grey shadings embed the columns (here representing nodes) belonging to the Caryophyllales and Cactaceae, respectively, associating to the similar shadings in Figs 2 and 3. Double vertical line delimits the fossil constrained node that is topologically the closes to the Cactaceae: nodes on the left of the double vertical line are constrained by fossils, nodes to the right of the double vertical line are not constrained by age. UCLN = Uncorrelated lognormal dating; NPRS = nonparametric rate smoothing; PLFB = Penalized likelihood with fossils-based rate smoothing; PLBP = penalized likelihood with branch-pruning rate smoothing; x = cross validation failed; not applicable (1) = applying this method requires at least one fixed and two constrained ages; not applicable (2) = applying this method requires at least 1 fixed age; pruned = excluded from the analysis; n.a.= not available; (a) = excluding *Cleistocactus icosagonus* and *Cleistocactus sepium*; (b) = excluding *Blossfeldia liliputana*; (c) = excluding *Cleistocactus icosagonus*. Column "mixing & converging" refers to the mixing and converging of the runs and is rated subjectively. Column "nb of runs" refers on how many runs in BEAST were needed to attain all of the ESS (effective size sample) values above at least 100.

Supplement S9. Comparison of topologies. Topologies yielded by different analytical methods are compared and monophyly and relationships of major clades are outlined.

Supplement S10. Variation of the optimal smoothing parameters selected through the cross-validation procedures for dsD: S10a shows PLFB cross-validation; S10b PLBP shows cross-validation.
Variation of the optimal smoothing parameters determined by applying large and small increments to a wide range of smoothing values; in one step or in few consequential steps in which, each time the range is limited. PLFB = penalized likelihood with fossil-based rate smoothing; PLBP = penalized likelihood with branch-pruning rate smoothing; λ = optimal rate smoothing; Ma = million years ago; CG = crown group.

Fig. 1 Strategies employed in this study

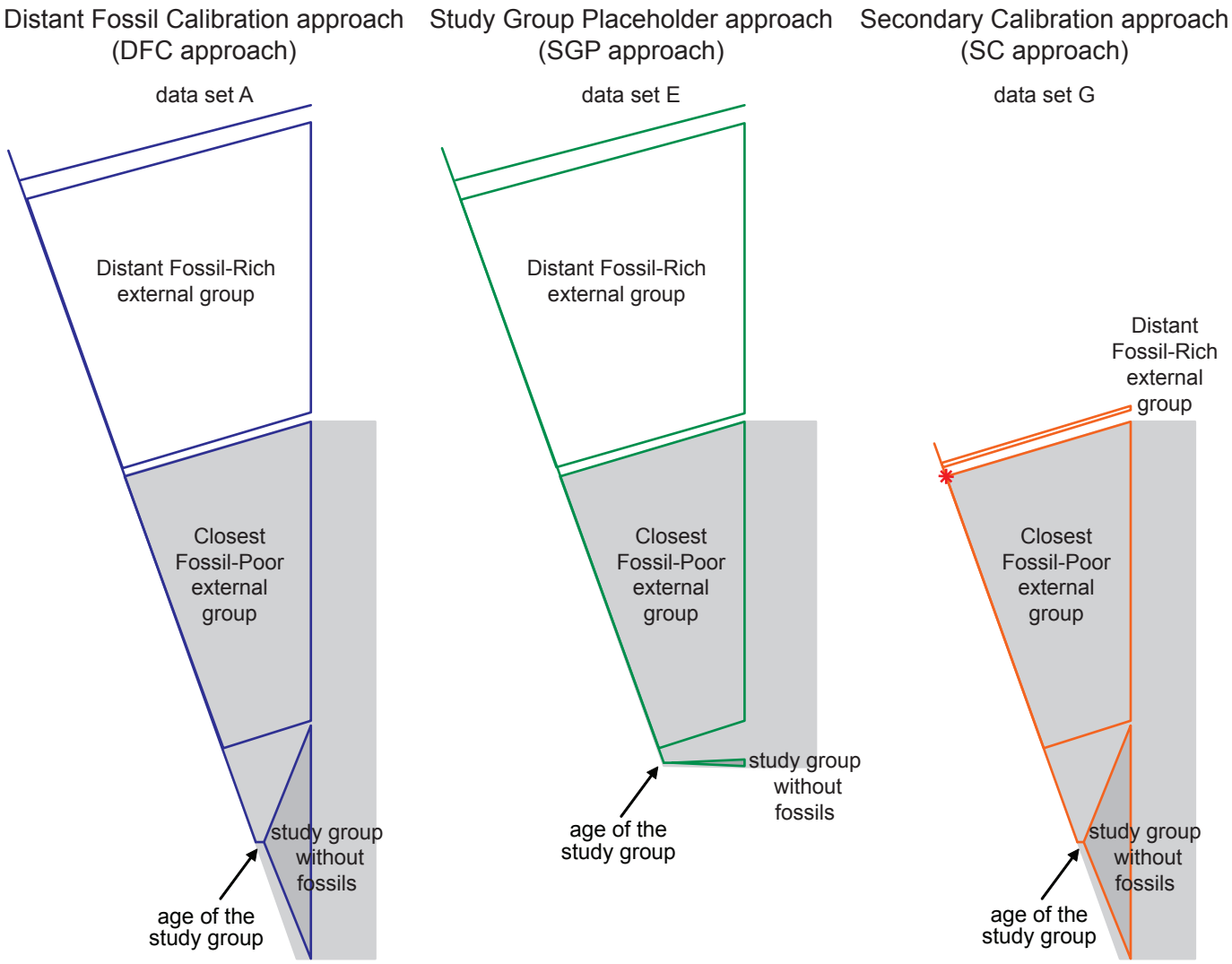


Fig. 2 Phylogram of a selected tree

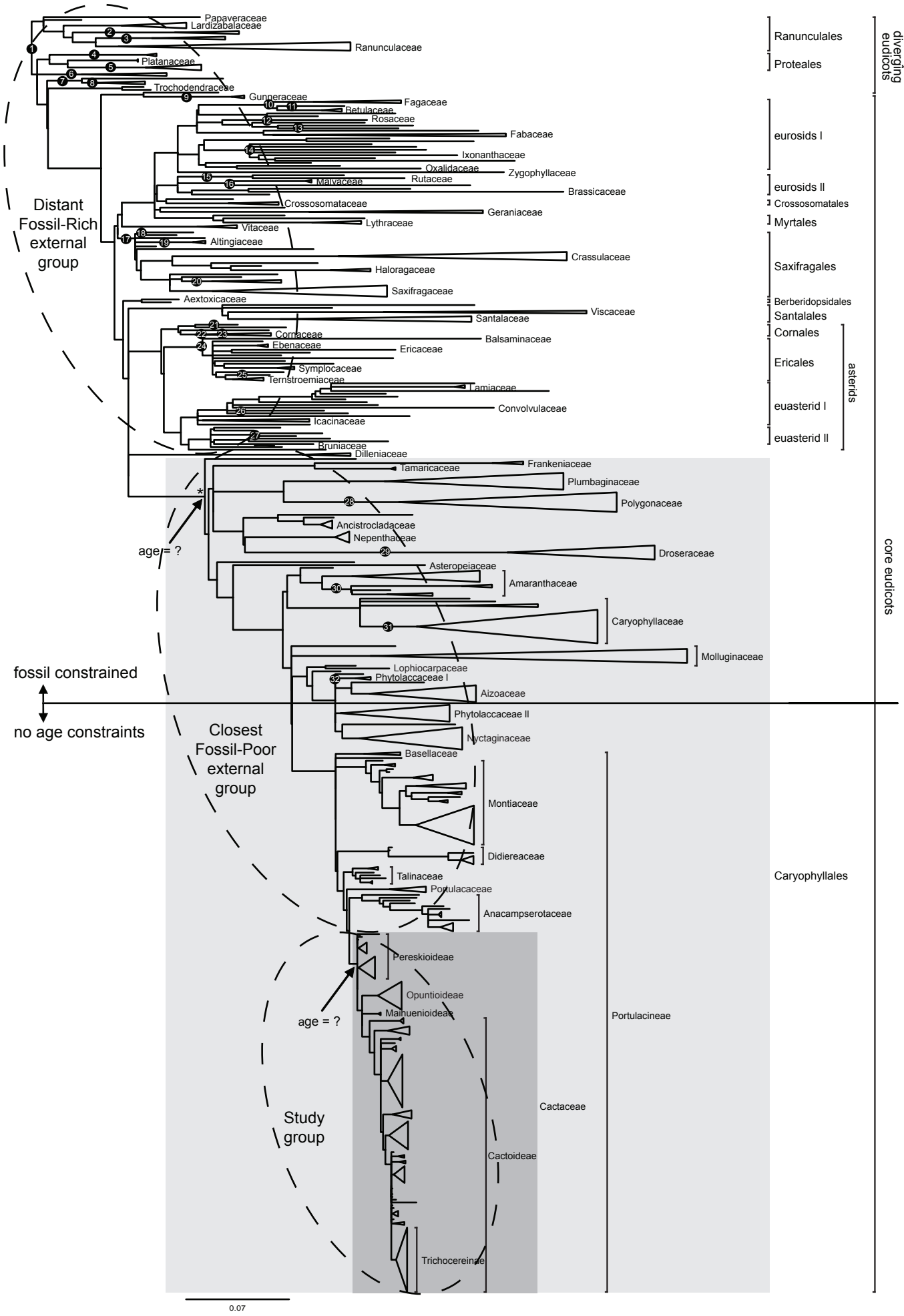


Fig. 3 BEAST Chronogram

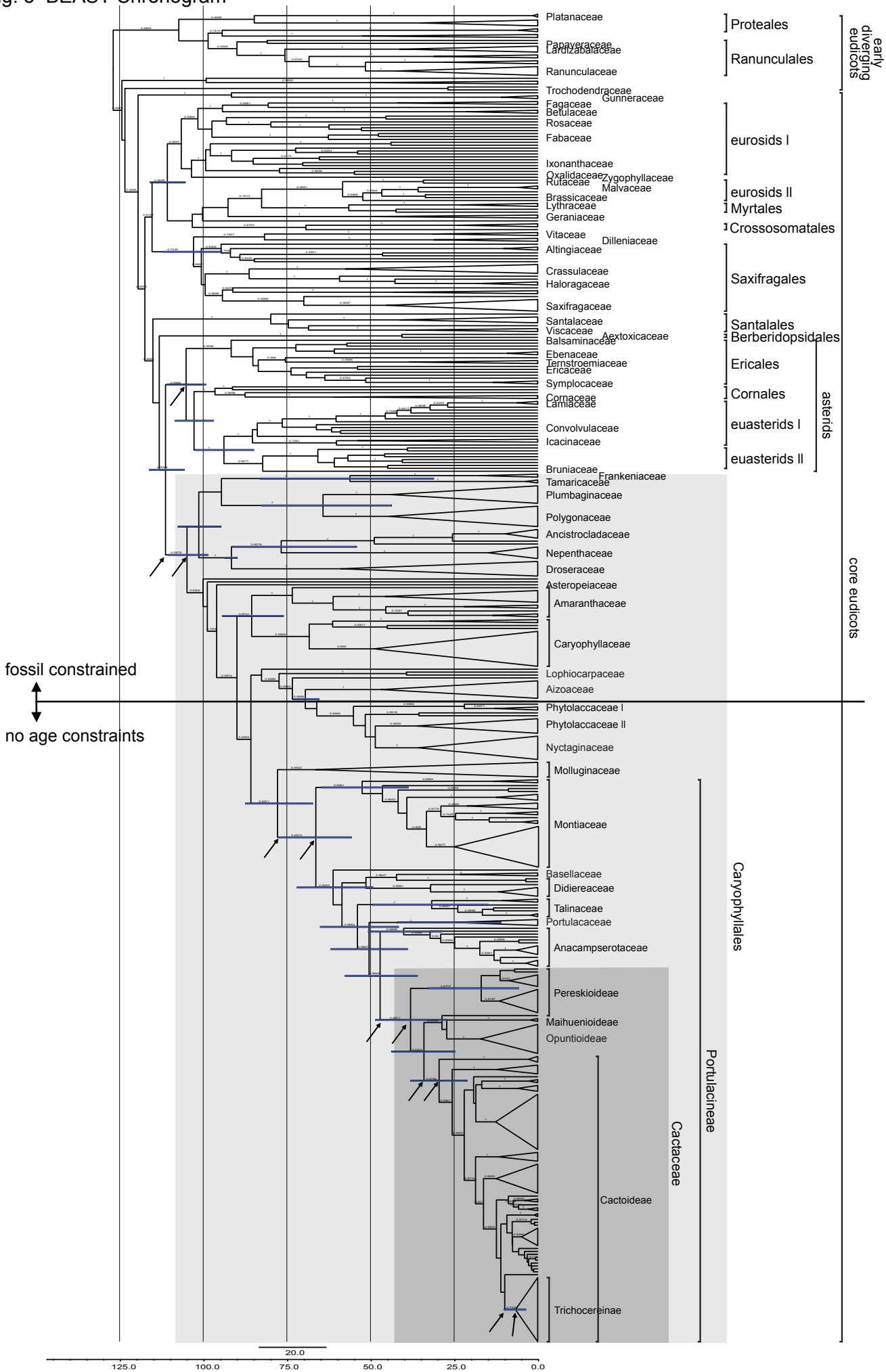


Fig. 4 Divergence ages

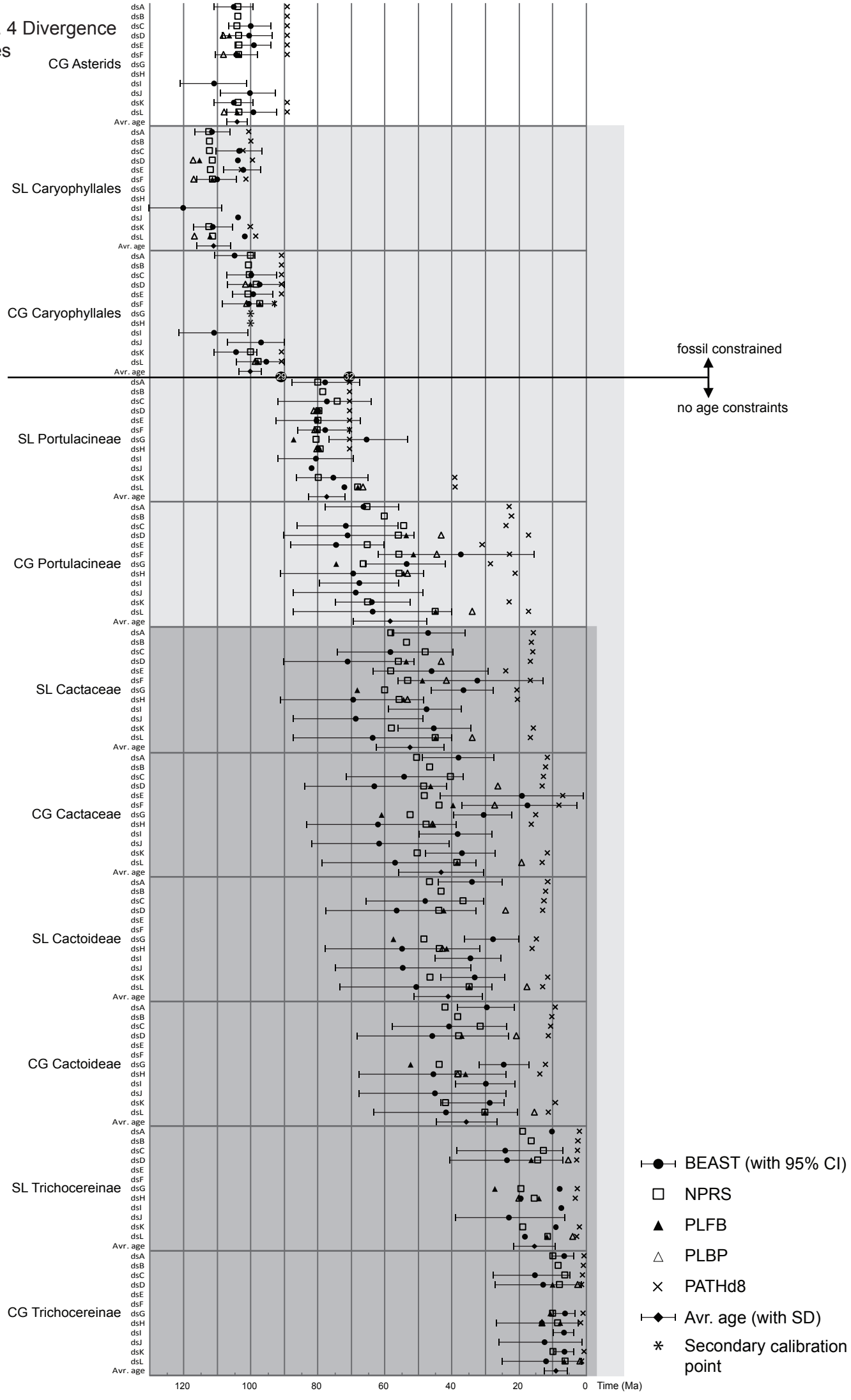


Fig. 5a UCLN - dense taxon sampling

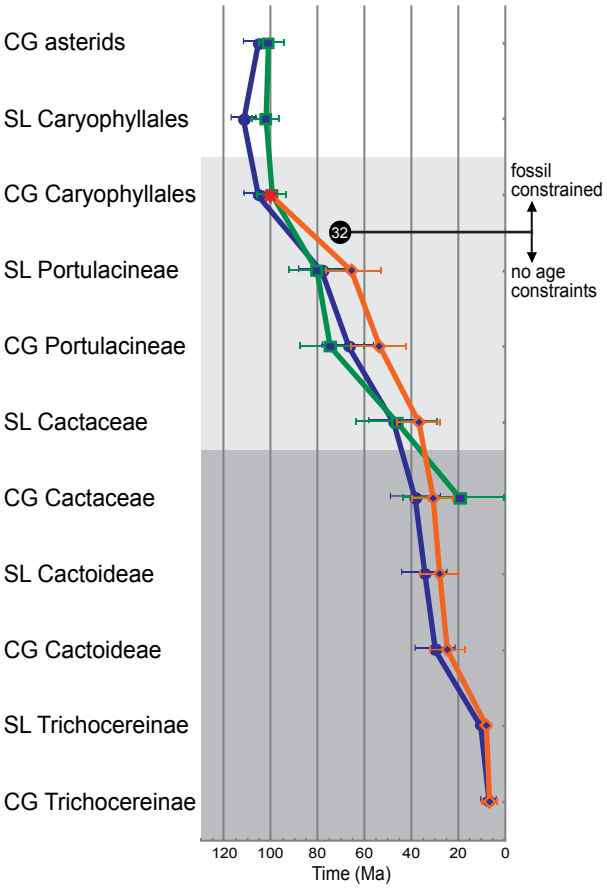


Fig. 5b UCLN - sparse taxon sampling

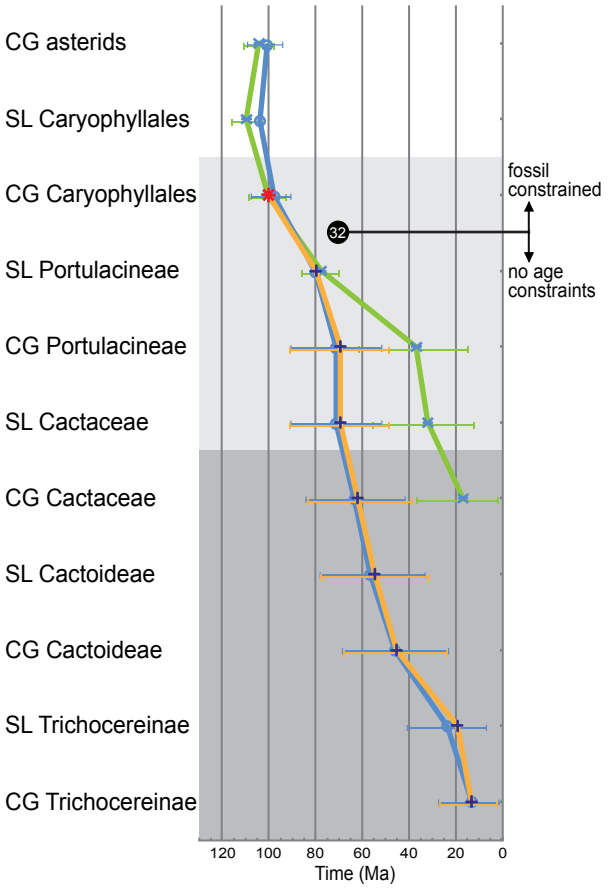


Fig. 5c NPRS - dense taxon sampling

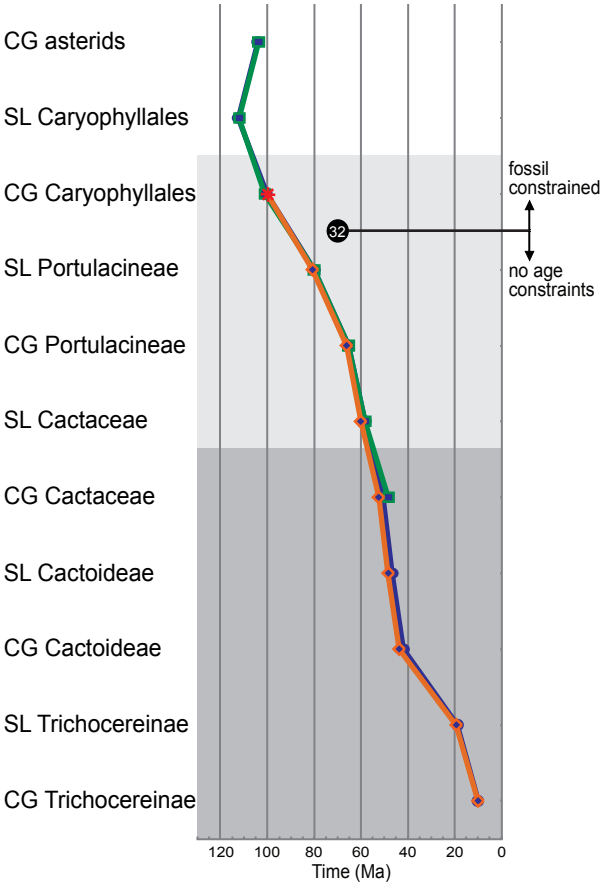
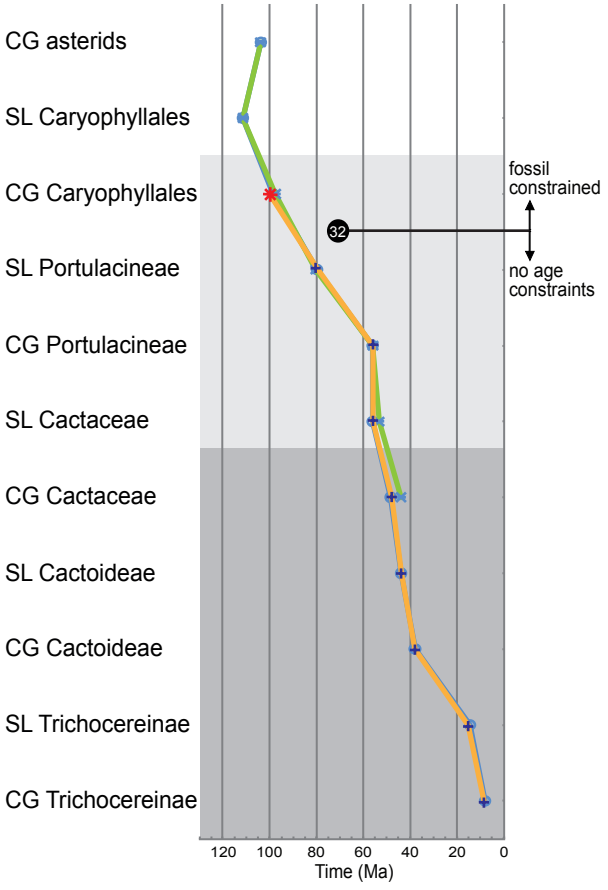


Fig. 5d NPRS - sparse taxon sampling



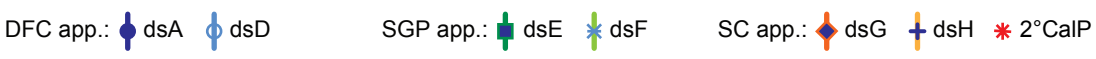


Fig. 5e PL - dense taxon sampling

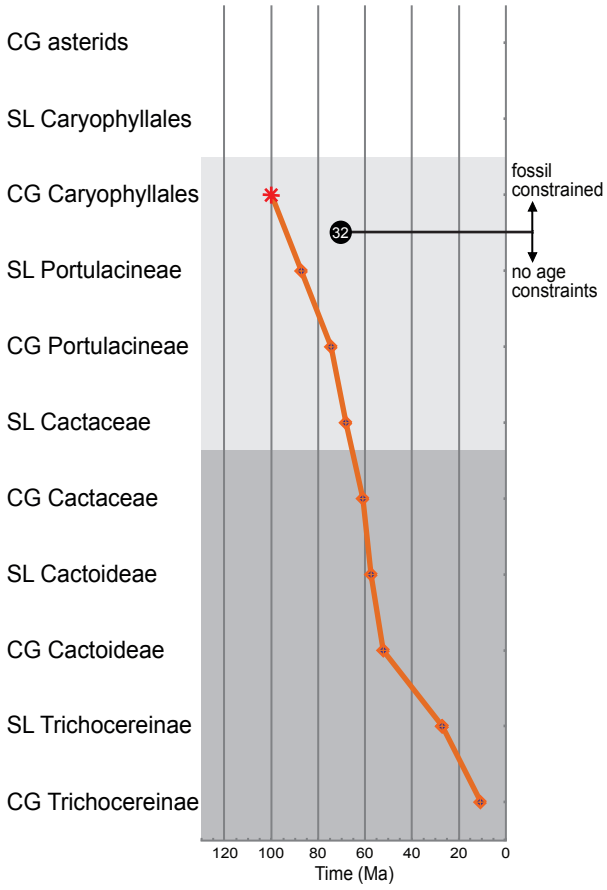


Fig. 5f PL - sparse taxon sampling

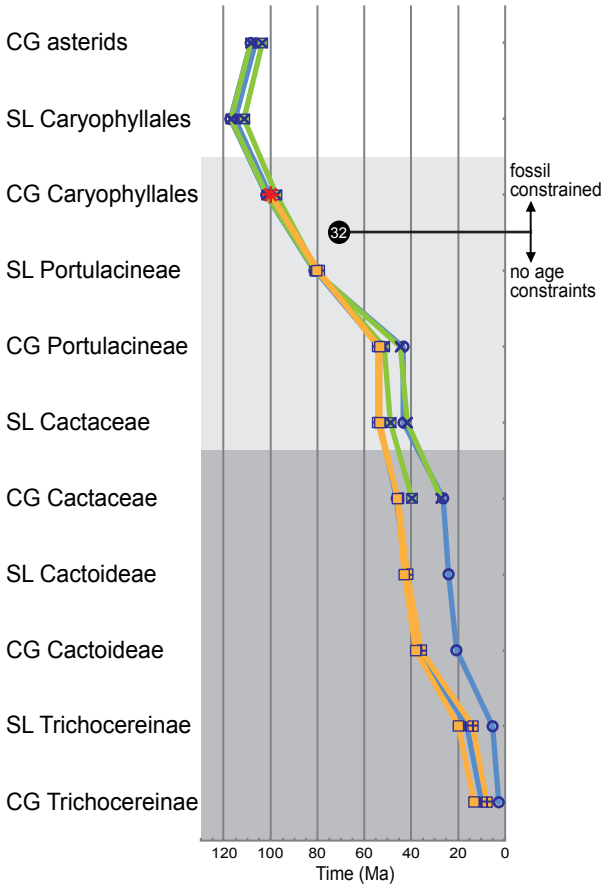


Fig. 5g PATHd8 - dense taxon sampling

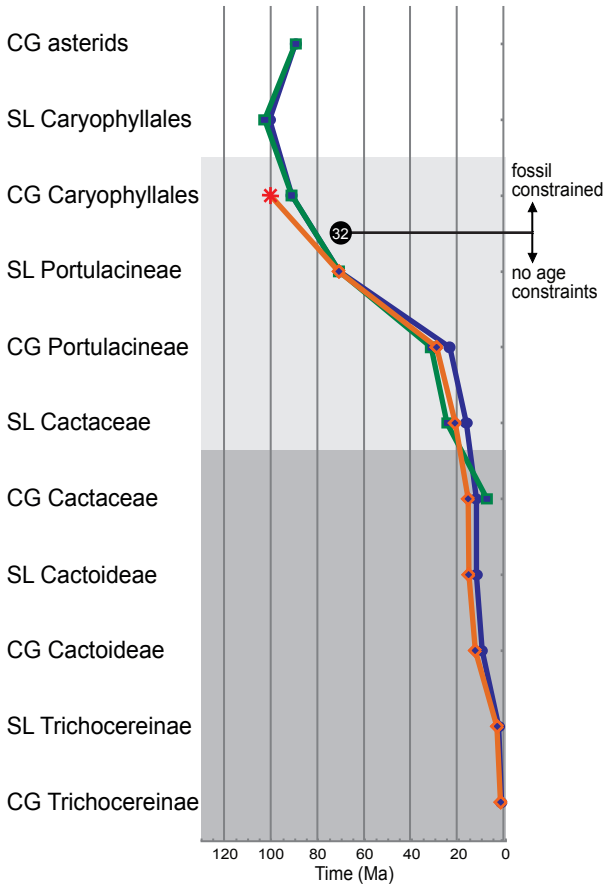
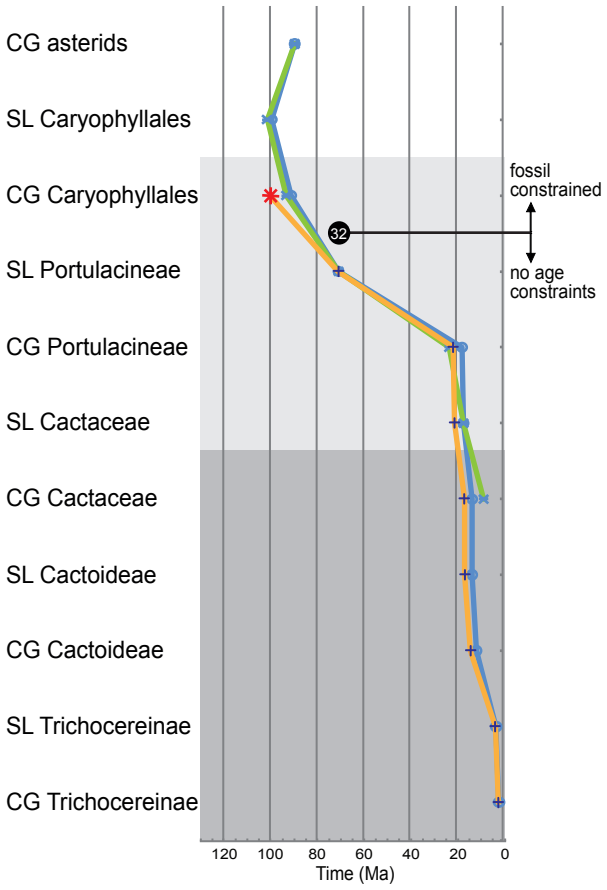


Fig. 5h PATHd8 - sparse taxon sampling



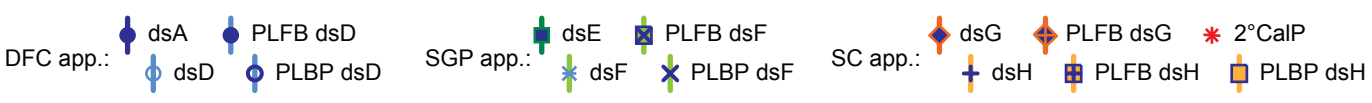


Fig. 6a UCLN - Distant Fossil Calibration approach

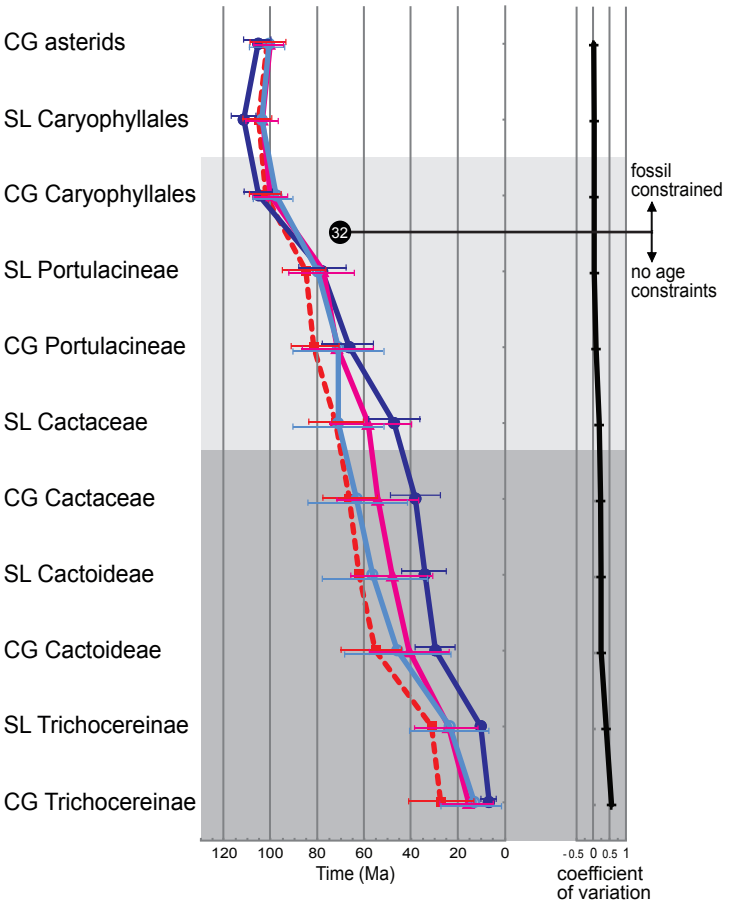


Fig. 6b NPRS - Distant Fossil Calibration approach

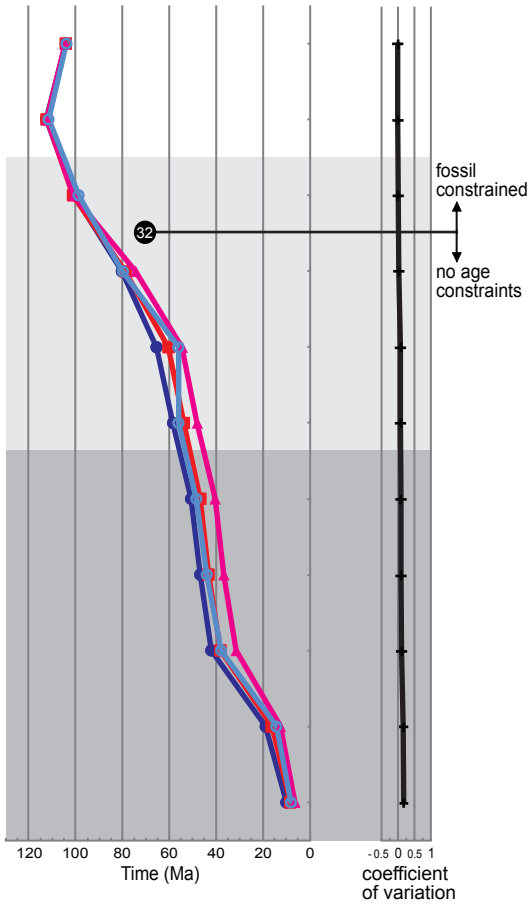


Fig. 6c PL - Distant Fossil Calibration approach

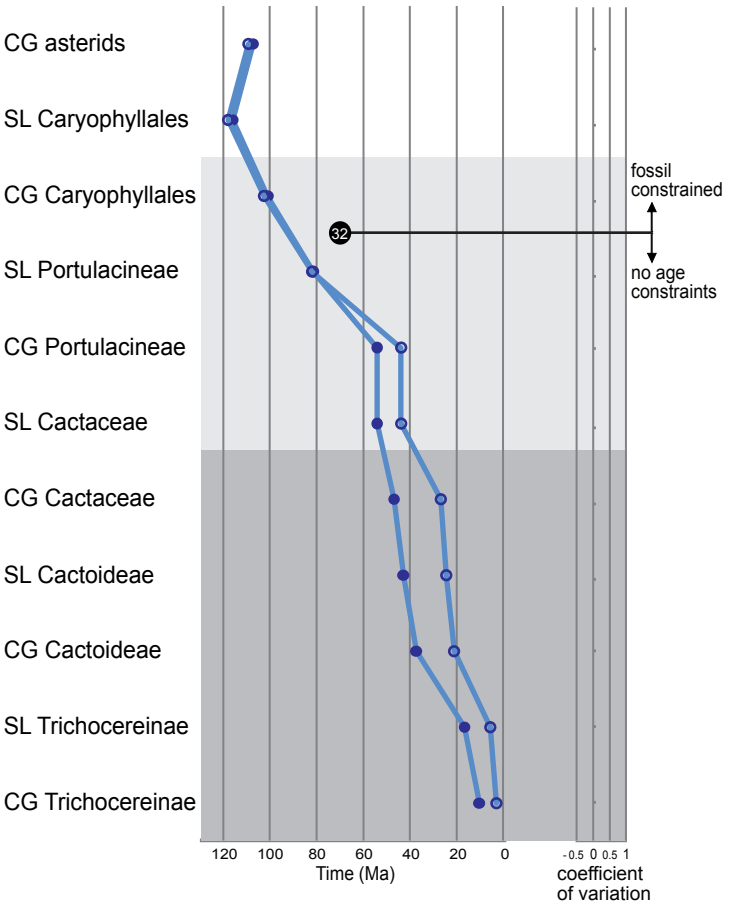
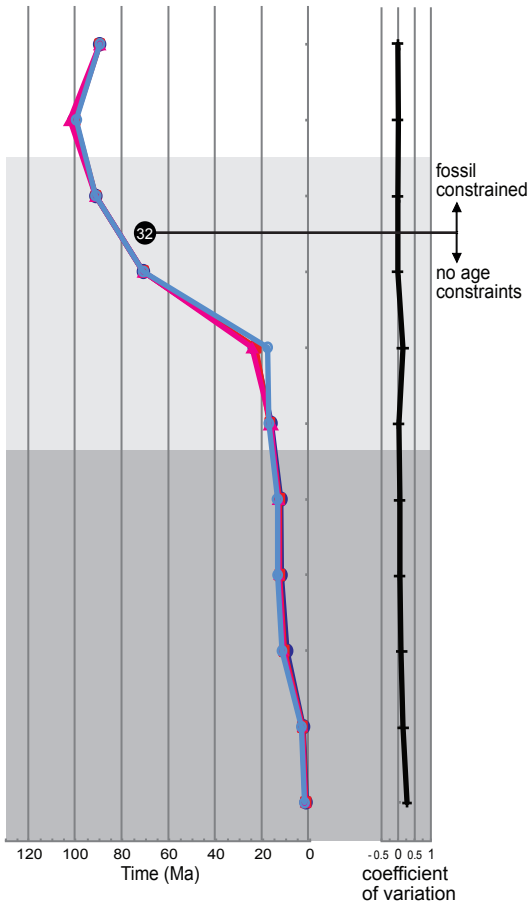


Fig. 6d PATHd8 - Distant Fossil Calibration approach



dsA dsB dsC dsD PLFB dsD PLBP dsD

Fig. 6e UCLN - restricted data sets

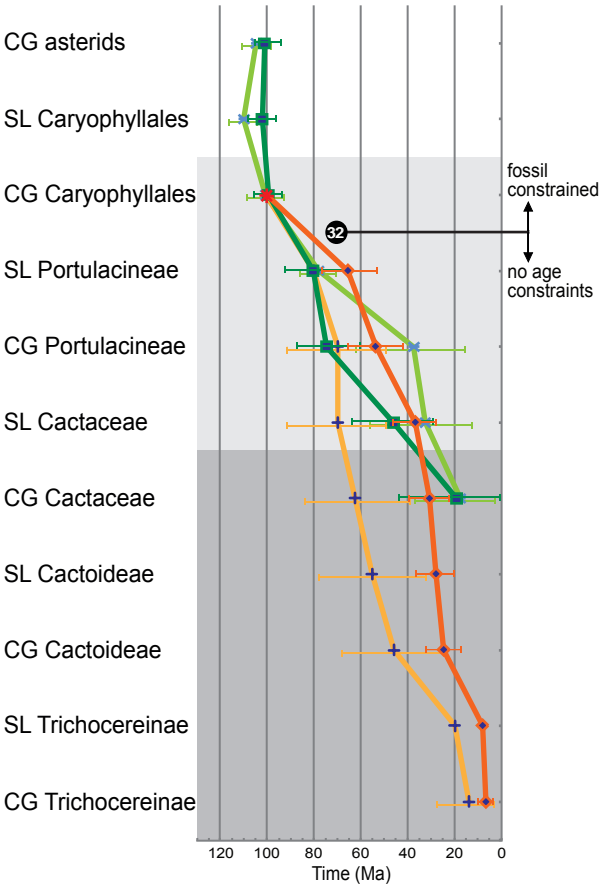


Fig. 6f NPRS - restricted data sets

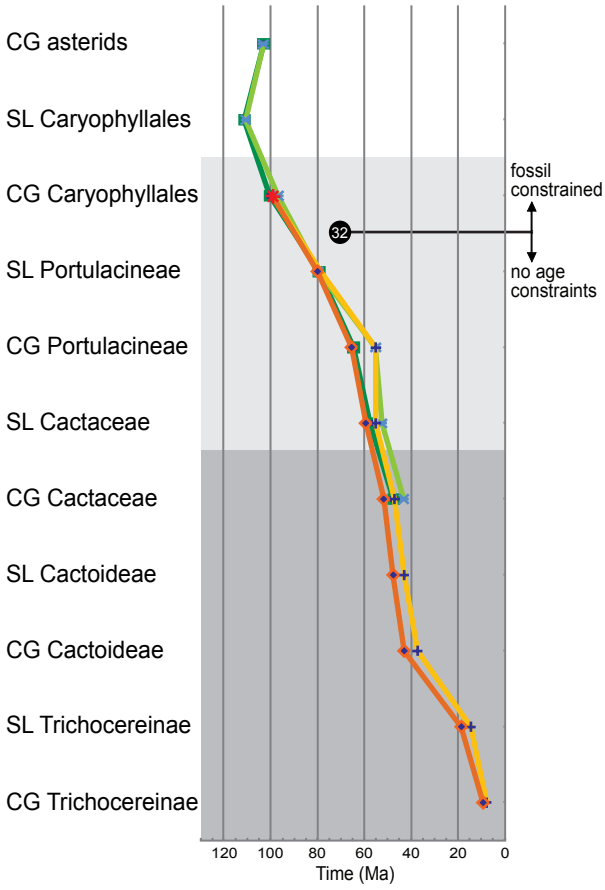


Fig. 6g PL - restricted data sets

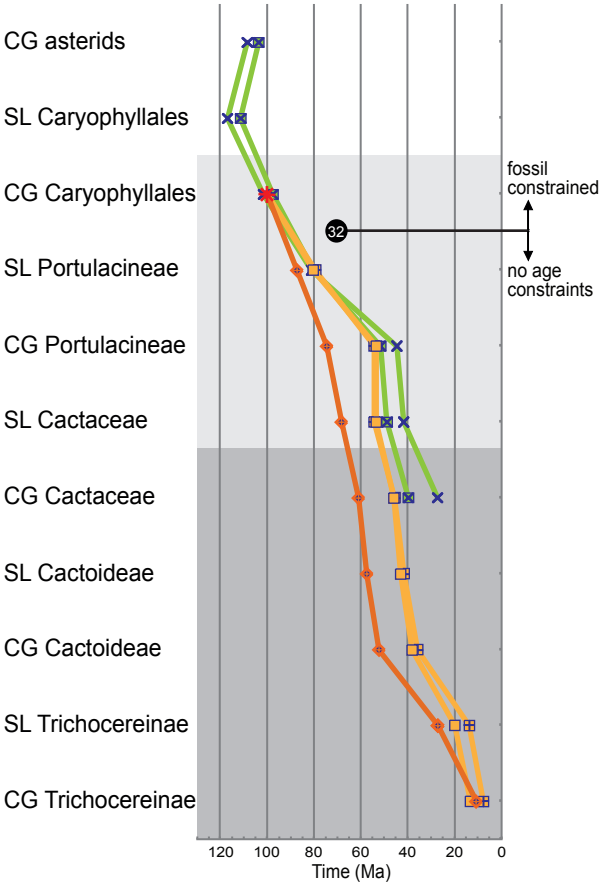


Fig. 6h PATHd8 - restricted data sets

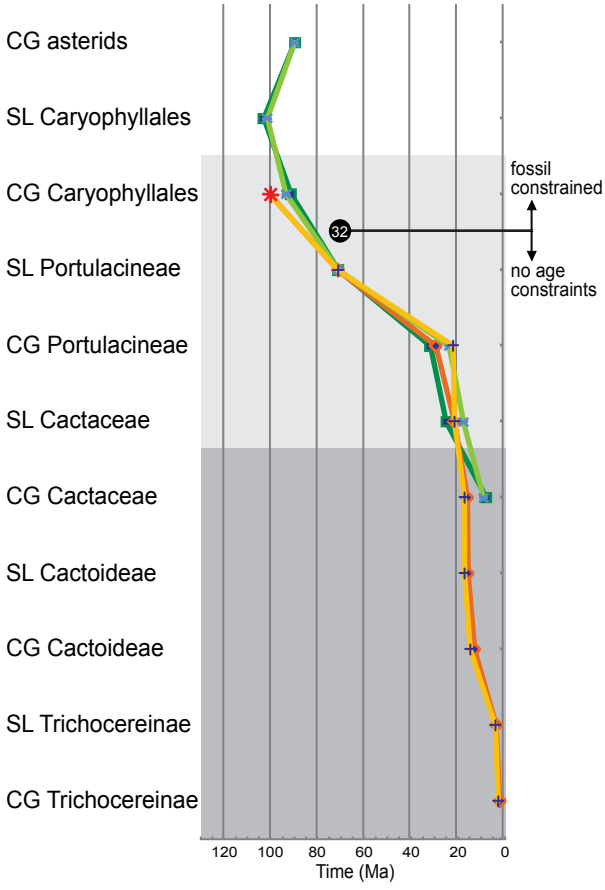
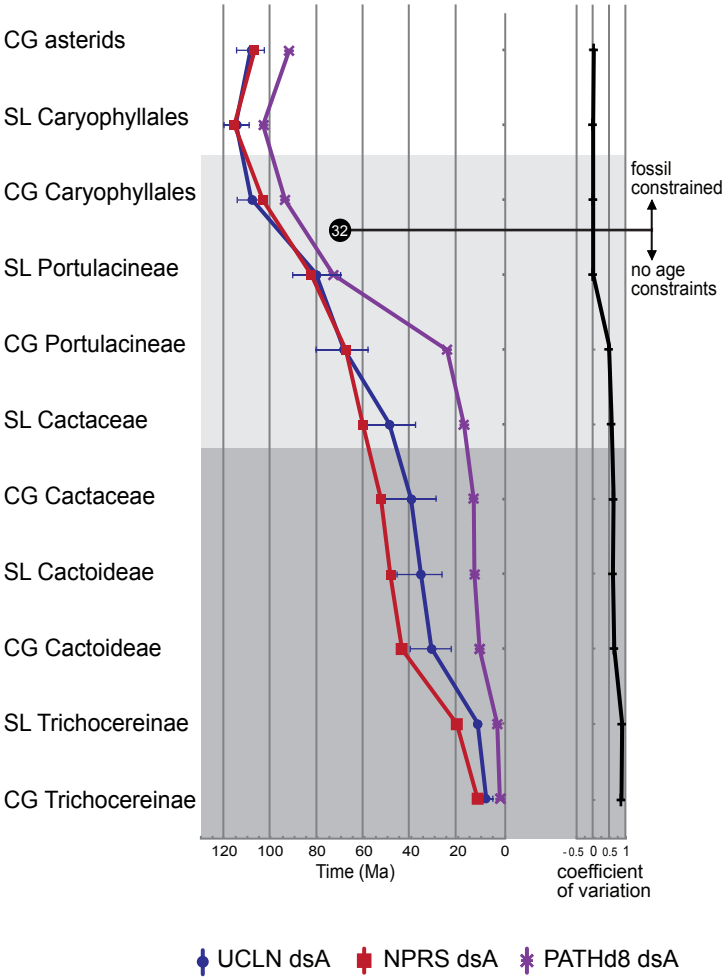
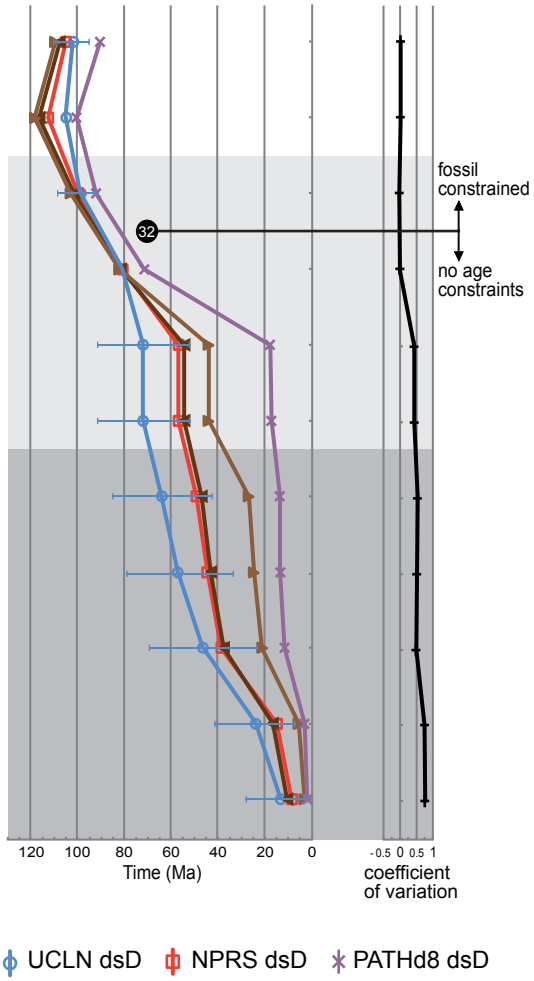


Fig. 7a dense taxon sampling



UCLN dsA NPRS dsA PATHd8 dsA

Fig. 7b sparse taxon sampling



UCLN dsD NPRS dsD PATHd8 dsD

PLFB dsD PLBP dsD

Fig. 8a UCLN - dense taxon sampling

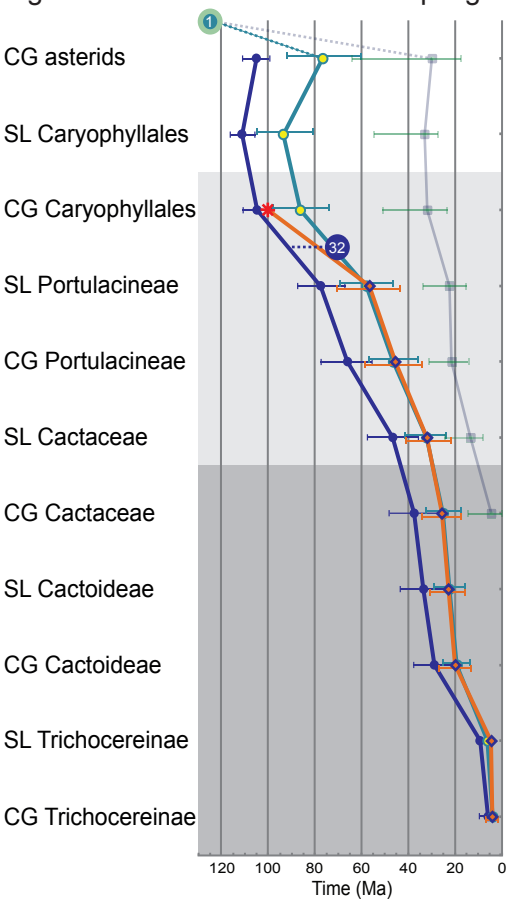


Fig. 8b UCLN - reduced taxon sampling

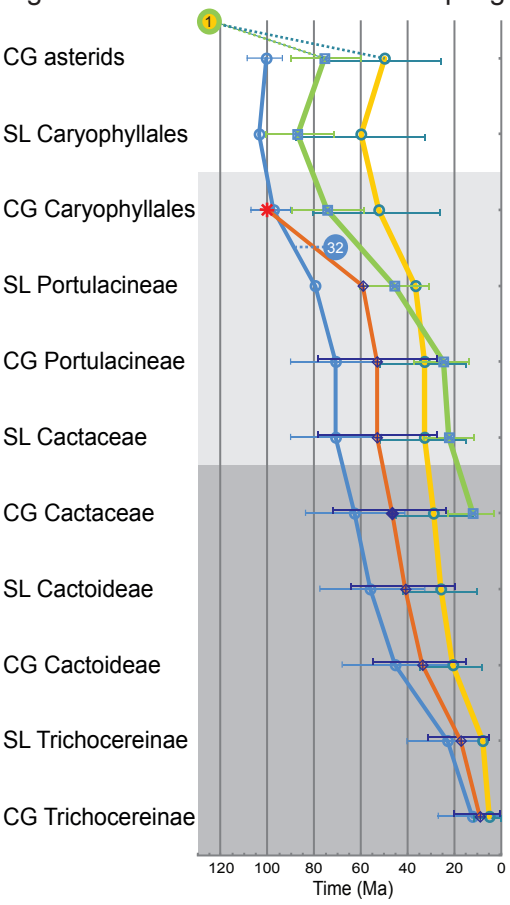


Fig. 8c NPRS - dense taxon sampling

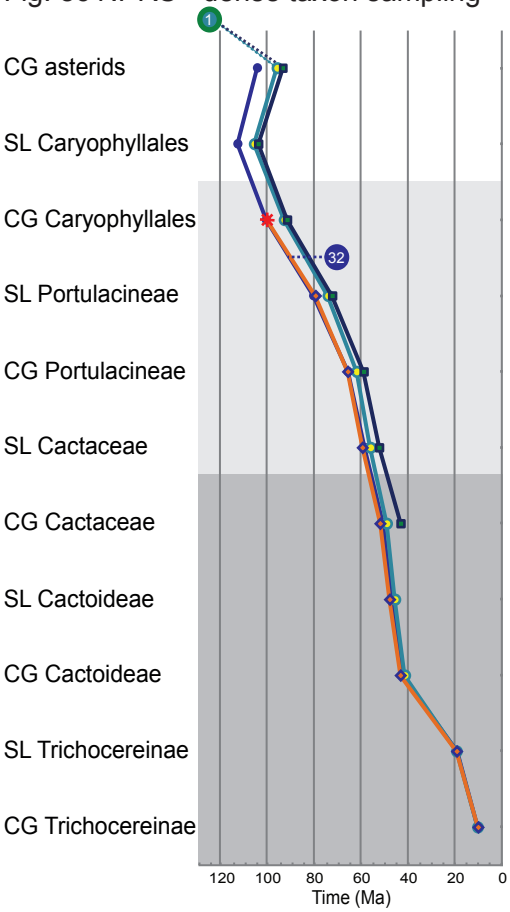


Fig. 8d NPRS - reduced taxon sampling

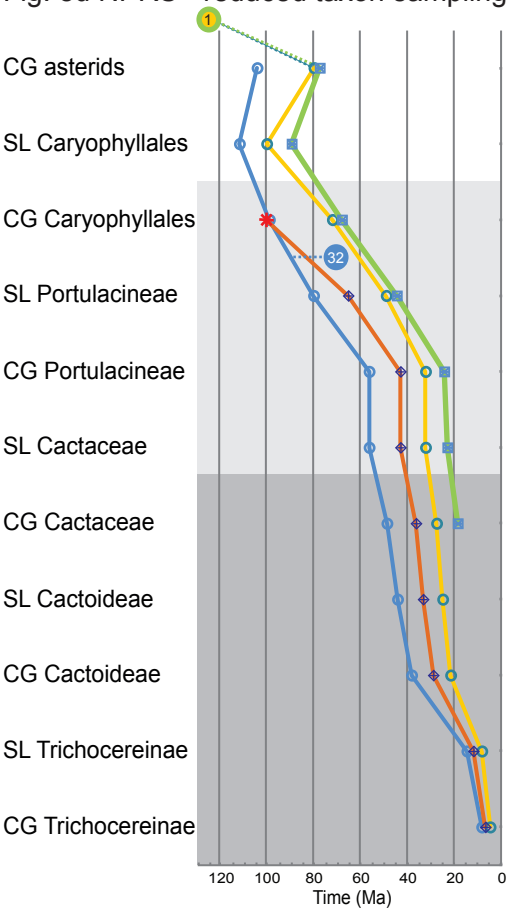


Fig. 8e PL - dense taxon sampling

PLFB: not applicable
(at least one fixed and two
constrained nodes are demanded)

PLBP: cross validation failed

Fig. 8f PL - reduced taxon sampling

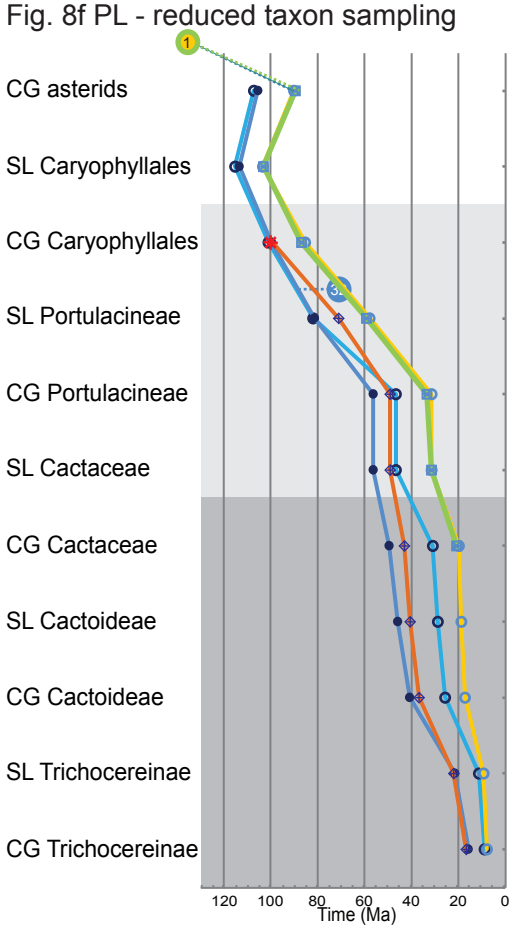


Fig. 8g PATHd8 - dense taxon sampling

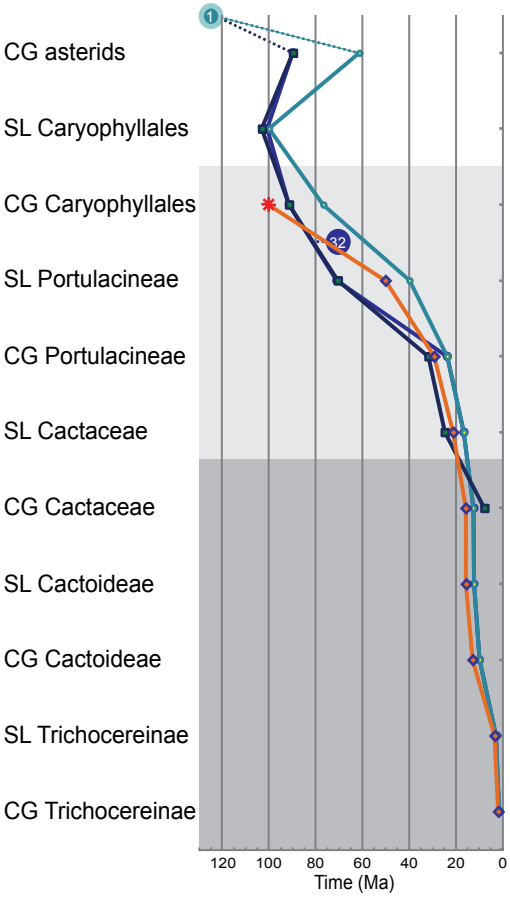
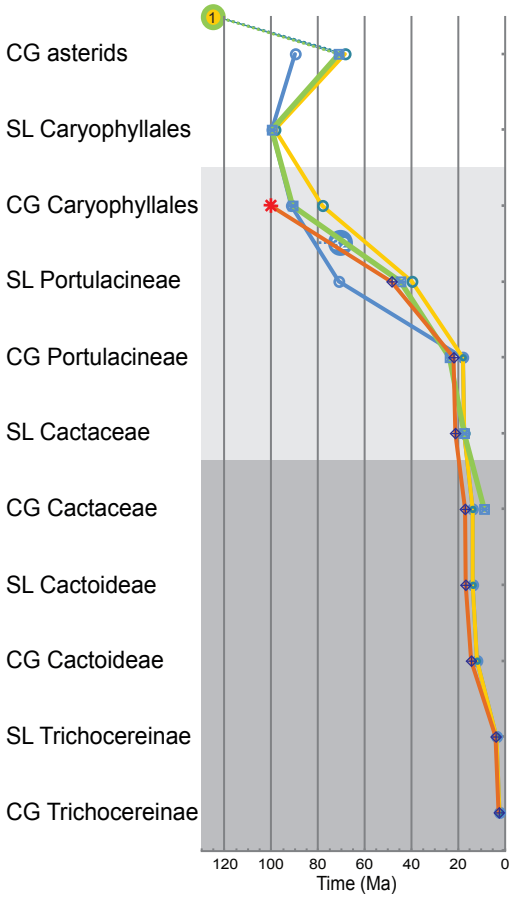


Fig. 8h PATHd8 - reduced taxon sampling



DFC app.: dsA, dsD, PLFB dsD, PLBP dsD, dsM, dsN

SGP app.: dsO, dsP

Single SC app.: 2'CalP, dsQ, dsR

Fig. 9 Divergence times estimated for the CG Cactaceae in penalized likelihood, when two different smoothing optimization criteria are applied

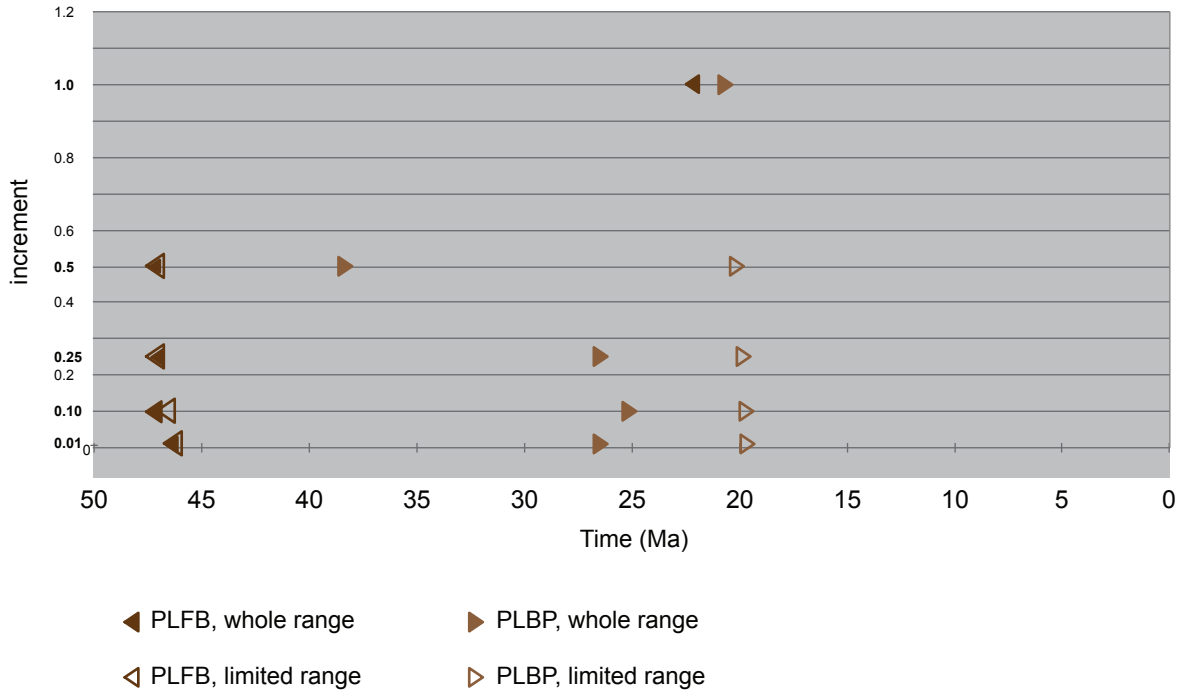


Fig. 10. Divergence times estimated for the CG Cactaceae in penalized likelihood when the same rate smoothing parameter is applied on different data sets

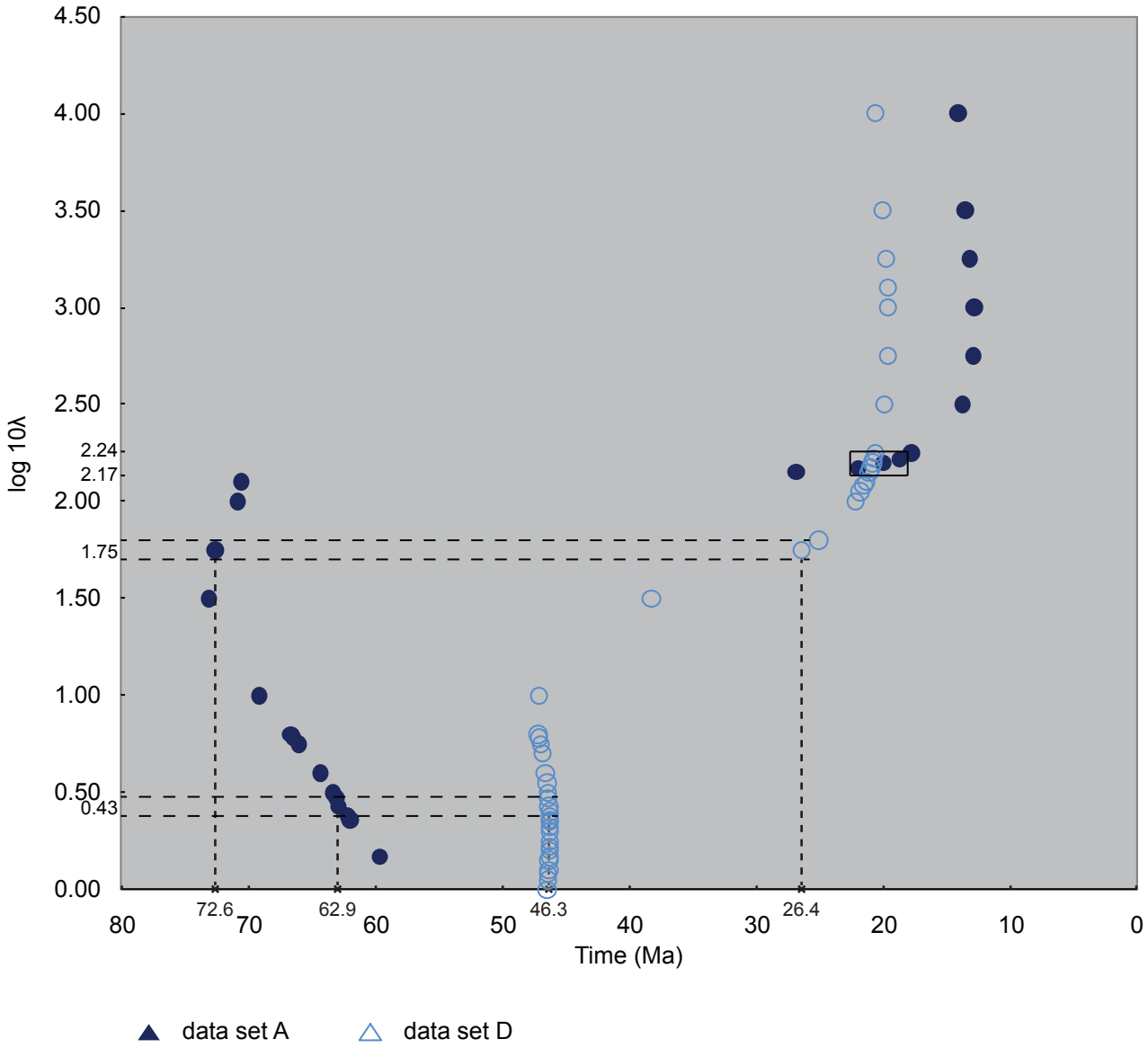


Table 1: Description of data sets

Approach	Taxon sampling strategy	Data set	Taxon sampling density	total number of taxa	number of taxa within: external groups: eudicots excluding Caryophyllales (Distant Fossil-Rich external group)	number of taxa within: external groups: Caryophyllales excluding Cactaceae (Closest Fossil-Poor external group)	number of taxa within: Caryophyllales: Cactaceae (study group)	Fossil sampling	total number of fossils (Fossil number in Fig. 2 and Appendix S1)	Description
Distant Fossil Calibration (DFC approach)	**Comprehensive**	**dsA**	Dense	**460**	161	170	129	**Comprehensive**	**32** (1-32)	**Reference data set** with a dense taxon sampling; comprehensive fossil sampling
		dsB	3/4 of the Dense	**345**	139	123	83		**32** (1-32)	Derived from data set A by sampling taxa sparsely to a 3/4 of a dense taxon sampling; comprehensive fossil sampling
		dsC	1/2 of the Dense	**230**	125	70	35		**32** (1-32)	Derived from data set A by sampling taxa sparsely to a 1/2 of a dense taxon sampling; comprehensive fossil sampling
		dsD	Sparse	**115**	70	20	25		**32** (1-32)	Derived from data set A by sampling taxa sparsely to a 1/4 of a dense taxon sampling; comprehensive fossil sampling; reference data set for following data sets: F, H, J, L and
Study Group Placeholder (SGP approach)	**Restricted** to external groups	**dsE**	Dense	**333**	161	170	2	**Comprehensive**	**32** (1-32)	Derived from data set A by restricting taxon sampling to external groups; comprehensive fossil sampling
		dsF	Sparse	**87**	70	15	2		**32** (1-32)	Derived from data set D by restricting taxon sampling to external groups; comprehensive fossil sampling
Secondary Calibration (SC approach)	**Restricted** to Caryophyllales	**dsG**	Dense	**300**	1	170	129	Fixed constraint **replaced**; fossils sampling **restricted** to 5 fossils from the Closest Foosil-Poor external group	**6** (2'CalP + 28-32)	Derived from data set A by restricting taxon and fossil sampling to Caryophyllales; a fixed age constraint of 100 Ma is added
		dsH	Sparse	**46**	1	20	25		**6** (2'CalP + 28-32)	Derived from data set D by restricting taxon and fossil sampling to Caryophyllales; a fixed age constraint of 100 Ma is added
excluding fixed (maximum) age constraint in DFC approach	**Comprehensive**	**dsI**	Dense	**460**	161	170	129	**Corroborated**	**31** (2-32)	Derived from data set A by excluding the fixed age
		dsJ	Sparse	**115**	70	20	25		**31** (2-32)	Derived from data set D by excluding the fixed age
excluding uncorroborated fossil calibrations in DFC approach	**Comprehensive**	**dsK**	Dense	**460**	161	170	129	**Corroborated**	**29** (1-29)	Derived from data set A by excluding 3 uncorroborated fossil constraints from the Closest Fossil-Poor external group
		dsL	Sparse	**115**	70	20	25		**29** (1-29)	Derived from data set D by excluding 3 uncorroborated fossil constraints from the Closest Fossil-Poor external group
single fossil calibration applied in DFC approach	**Comprehensive**	**dsM**	Dense	**460**	161	170	129	**Corroborated**	**1** (1)	Derived from data set A by reducing the fossil sampling to a single, distant and fixed fossil constraint
		dsN	Sparse	**115**	70	20	25		**1** (1)	Derived from data set D by reducing the fossil sampling to a single, distant and fixed fossil constraint
single fossil calibration applied in (SGP approach)	**Restricted** to external groups	**dsO**	Dense	**333**	161	170	2	**Corroborated**	**1** (1)	Derived from data set A by restricting taxon sampling to external groups and a fossil sampling to a single, distant and fixed fossil
		dsP	Sparse	**87**	70	15	2		**1** (1)	Derived from data set D by restricting taxon sampling to external groups and a fossil sampling to a single, distant and fixed fossil
single Secondary Calibration Point	**Restricted** to Caryophyllales	**dsQ**	Dense	**300**	1	170	129	Fixed constraint **replaced**;	**1** (2'CalP)	Derived from data set A by restricting taxon sampling to Caryophyllales and fixing the age of CG Caryophyllales to 100 Ma
		dsR	Sparse	**46**	1	20	25		**1** (2'CalP)	Derived from data set D by restricting taxon sampling to Caryophyllales and fixing the age of CG Caryophyllales to 100 Ma

Supplement S1: Fossil constraints

Fossil nb.	Node	Age (Ma)	Fossil	Source of age	Original reference
1	CG Eudicots	125	*Sinocarpus decussatus*	Magallón and Castillo, 2009	Leng and Friis, 2003
2	SL Menispermaceae	65.5	Menispermaceae endocarps	Magallón and Castillo, 2009	Friis et al., 2006
3	SL Berberidaceae	33.9	Mahonia leaves	Magallón and Castillo, 2009	Manchester, 1999
4	SL Nelumbonaceae	106	*Nelumbites extenuinervis*	Anderson et al., 2005	Vakhrameev, 1952
5	SL Proteaceae	85	*Beaupreopsis and Macadamia*	Anderson et al., 2005	Dettmann and Jarzen, 1998
6	SL Sabiaceae	98	*Insitiocarpus moravicus*	Anderson et al., 2005	Knobloch and Mai, 1986
7	SL Buxales	99.6	*Spanomera marylandensis*	Magallón and Castillo, 2009	Drinnan et al., 1991
8	SL Buxaceae	94	*Spanomera marylandensis*	Anderson et al., 2005	Drinnan et al., 1991
9	SL Gunneraceae	91	*Retitricolpites microreticulatus*	Anderson et al., 2005	Jarzen and Dettmann, 1989
10	CG Fagales	93.5	Normapolles pollen	Magallón and Castillo, 2009	Batten, 1981, Kedves, 1989, Pacltová, 1966
11	SL Juglandaceae	83.5	*Antiquocarya spp., Caryanthus spp., Manningia crassa*	Magallón and Castillo, 2009	Friis, 1983
12	SL Rhamnaceae	48.6	*Pailurus sp.*	Magallón and Castillo, 2009	Manchester, 1999
13	SL Ulmaceae	55.8	Extinct Ulmaceae	Magallón and Castillo, 2009	Manchester, 1999
14	CG Malpighiales	89.3	*Paleoclusia chevalieri*	Magallón and Castillo, 2009	Crepet and Nixon, 1998
15	SL Sapindales	55.8	*Acer sp., Dipteronia sp. Koelreutia sp.*	Magallón and Castillo, 2009	Manchester, 1999
16	CG Malvales	33.9	*Craigia sp., Tilia sp.*	Magallón and Castillo, 2009	Manchester, 1999
17	SL Saxifragales	91	*Divisestylus*	Anderson et al., 2005	Hermsen et al., 2003
18	CG Hamamelidaceae	83.5	*Allonia decandra, Androdecidua endressii*	Magallón and Castillo, 2009	Magallon-Puebla et al., 1996, Magallon et al 2001
19	SL Altingiaceae	89.3	*Microaltingia apocarpelata*	Magallón and Castillo, 2009	Zhou et al., 2001
20	SL Iteaceae	89.3	*Divisestylus brevistamineus*	Magallón and Castillo, 2009	Hermsen et al., 2003
21	SL Hydrangeaceae	89.3	*Tylerianthus crossmaniensis*	Magallón and Castillo, 2009	Gandolfo, 1998
22	SL Cornaceae	87	fruit	Anderson et al., 2005	Takahashi et al., 2002
23	CG Cornaceae	55.8	*Cornus spp.*	Magallón and Castillo, 2009	Mai, 1993 (Manchester 1999)
24	CG Ericales	89.3	*Paleoenkianthus sayrevillensis*	Magallón and Castillo, 2009	Nixon and Crepet, 1993
25	SL Pentaphylacaceae	65.5	*Eurya sp.*	Magallón and Castillo, 2009	Knobloch and Mai, 1986,

Supplement S1: Fossil constraints

					Bremer et al.., 2004
26	SL Vahliaceae	83.5	*Scandianthus spp.*	Magallón and Castillo, 2009	Friis and Skarby, 1982
27	CG Apiales	37.2	*Torciellia bonensii*	Magallón and Castillo, 2009	Manchester, 1999
28	SL Polygonaceae	5.33	Polygonaceae fruits	Magallón and Castillo, 2009	Gregor, 1982, Friis, 1985, van Der Burgh, 1987, Dorofeev, 1988
29	SL Droseraceae	91	*Palaeoaldrovanda splendens*	Anderson et al., 2005	Knobloch and Mai, 1986
30	CG Chenopodioideae / Amaranthaceae	55.8	*Chenopodipollis multiplex*	Kadereit et al., 2005	Kadereit, G.et al_. 2005.
31	SL higher Caryophyllaceae *	33.9	*Caryophylloflora paleogenica*	Jordan and Macphail, 2003	Jordan and Macphail, 2003
32	SL Phytolaccaceae s.str.	70.6	*Coahuilacarpon phytolaccoides*	Magallón and Castillo, 2009	Cevallos-Ferriz et al., 2008

Supplement S2: Published ages for Cactaceae and subclades

Published ages:	method	CG asterids	SL Caryophyllales	CG Caryophyllales	SL Portulacineae	CG Portulacineae	SL Cactaceae	CG Cactaceae	SL Cactoideae	CG Cactoideae	SL Trichocereinae	CG Trichocereinae
current paper	UCLN dsA (CI)	105.1 (99.5 - 111.1)	111.3 (105.9 - 116.4)	104.8 (98.7 - 110.7)	77.8 (67.5 - 87.6)	66.3 (55.9 - 77.7)	47.1 (36.2 - 57.8)	38.0 (27.4 - 48.7)	34.0 (24.9 - 43.9)	29.5 (21.3 - 38.2)	10.1	6.6 (3.8 - 10.3)
current paper	NPRS dsA	104.0	112.3	100.2	80.1	65.4	58.3	50.6	46.8	42.0	18.9	10.1
current paper	PLFB dsD	106.4	115.0	100.2	80.6	53.6	53.6	46.3	42.4	36.9	16.3	10.1
current paper: average age (Max - min)	UCLN, NPRS, PLFB, PLBP	104.2 (111.0 - 99.1)	110.8 (119.9 - 101.5)	100.2 (111.0 - 95.4)	77.3 (87.1 - 65.4)	58.4 (74.5 -34.1)	52.5 (71.1 - 32.5)	43.2 (63.1 - 17.5)	41.1 (57.4 - 17.8)	35.6 (52.2 - 15.4)	15.3 (27.1 - 4.0)	9.1 (15.4 - 2.1)
Arakaki et al. 2011	multidivtime	-	ca 95	ca 75.8	53.3 (+- 2.1)	44.9 (+- 2.7)	35.0 (+- 2.6)	28.6 (+-1.9)	24.4 (+-1.0)	21.8 (1.7)	6.5 (+- 2.0)	5.3 (+- 1.9)
Ocampo et al. 2011	UCLN (95% CI)	-	-	-	18.8 (6.7 - 33.7)	-	13.9 (4.9 - 26.5)	10 (3.1 - 19.1)	-	-	-	-
Bell et al. 2010	UCLN# (95% CI)	106-114 (100-122)	111-121 (104-129)	99-106 (91-115)	-	-	21-22 (11-33)	-	-	-	-	-
Martínez-Millán 2010	fossil record	89.3	-	-	-	-	-	-	-	-	-	-
Moore et al. 2010	UCLN (95% CI)	84 (80-89)	100 (95-104)	67 (63-71)	-	-	-	-	-	-	-	-
Smith et al. 2010	UCLN	-	-	83.5	-	-	-	-	-	-	-	-
Magallon&Castillo,2009	PLFB (CI)	-	111.33 (111.18-111.48)	94.53 (94.45 -94.6)	-	-	-	-	-	-	-	-
Soltis et al. 2008	NPRS*; PATHd8*	91-108; 69-80	107-117; 72-86	-	-	-	-	-	-	-	-	-
Anderson et al. 2005	PL; NPRS	109.0	114; 116	99; 102	-	-	-	-	-	-	-	-
Bremer et al. 2004	PL, NPRS, strict molecular clock of Langley and Fitch	128 [96-156]	-	-	-	-	-	-	-	-	-	-
Crepet et al 2004	fossil record	90.0	-	-	-	-	-	-	-	-	-	-
Landrum 2002	morphology	-	-	-	-	-	>40	-	-	-	-	-
Wikström et al. 2001	NPRS	107-117 (+-5)	104-111 (+-4)	83-90 (+-4)	-	-	11-18 (+-2)	-	-	-	-	-
Hershkovitz&Zimmer 1997	other "	-	-	-	-	-	~30	-	-	-	-	-
Gibson & Nobel 1986	geological events	-	-	-	-	-	~100	-	-	-	-	-

Legend:

SL = stem lineage
CG = crown group

UCLN = an uncorrelated lognormal relaxed-clock model, with confidence interval values in brackets
NPRS = nonparametric rate smoothing method
PLFB = penalized likelihood with fossils-based rate smoothing
PLBP = penalized likelihood with branch-pruning rate smoothing
PL = penalized likelihood (the smoothing method is not specified)
ds= data set

Average age = average value excluding PATHd8, UCLN values for "B", and data sets "I" and "J"
SD = standard deviation
[ ] = values of the six analyses
CI = confidence interval
UCLN# = values from multiple-fossil analyses that treated all fossils as exponential or lognormal priors are shown together
NPRS*; PATHd8* = range between the lowest and the highest estimate of analyses with a single fixed and multiple calibration analyses is shown
other" = comparing ITS homogeneity with different lineages of angiosperms

Supplement S3

Group	Family	Species	GenBank accession ID	Taxon code
outgroup	Ceratophyllaceae	*Ceratophyllum demersum*	EF614270	Cer_demGnBk
outgroup	Ceratophyllaceae	*Ceratophyllum echinatum*	AY335986	Cer_echGnBk
outgroup	Ceratophyllaceae	*Ceratophyllum submersum*	DQ401361	Cer_subGnBk
early diverging eudicots	Lardizabalaceae	*Akebia quinata*	AF542587	Ake_quiGnBk
early diverging eudicots	Buxaceae	*Buxus sempervirens*	AF543728	Bux_semGnBk
early diverging eudicots	Berberidaceae	*Caulophyllum thalictroides*	AB069831	Cau_thaGnBk
early diverging eudicots	Papaveraceae	*Dicentra eximia*	DQ182345	Dic_exiGnBk
early diverging eudicots	Didymelaceae	*Didymeles perrieri*	DQ401354	Did_perGnBk
early diverging eudicots	Proteaceae	*Embothrium coccineum*	AM396515	Emb_cocGnBk
early diverging eudicots	Eupteleaceae	*Euptelea pleiosperma*	AM396510	Eup_pleGnBk
early diverging eudicots	Ranunculaceae	*Glaucidium palmatum*	AB069850	Gla_palGnBk
early diverging eudicots	Proteaceae	*Grevillea banksii*	AF542583	Gre_banGnBk
early diverging eudicots	Ranunculaceae	*Hydrastis canadensis*	AB069849	Hyd_canGnBk
early diverging eudicots	Sabiaceae	*Meliosma cuneifolia*	AM396513	Mel_cunGnBk
early diverging eudicots	Berberidaceae	*Nandina domestica*	AB069830	Nan_domGnBk
early diverging eudicots	Nelumbonaceae	*Nelumbo lutea*	EU642710	Nel_lutGnBk
early diverging eudicots	Nelumbonaceae	*Nelumbo nucifera*	AM396514	Nel_nucGnBk
early diverging eudicots	Buxaceae	*Pachysandra terminalis*	AF542581	Pac_terGnBk
early diverging eudicots	Platanaceae	*Platanus occidentalis*	EU642711	Pla_occGnBk
early diverging eudicots	Platanaceae	*Platanus racemosa*	EU169656	Pla_racGnBk
early diverging eudicots	Menispermaceae	*Pericampylus glaucus*	EF143869	Prc_glaGnBk
early diverging eudicots	Ranunculaceae	*Ranunculus trichophyllus*	AY954133	Ran_triGnBk
early diverging eudicots	Proteaceae	*Roupala montana*	EU642684	Rou_monGnBk
early diverging eudicots	Sabiaceae	*Sabia sp.*	DQ401352	Sab_sp_GnBk
early diverging eudicots	Lardizabalaceae	*Sinofranchetia chinensis*	EF143879	Sin_chiGnBk
early diverging eudicots	Lardizabalaceae	*Sargentodoxa cuneata*	DQ401351	Srg_cunGnBk
early diverging eudicots	Tetracentraceae	*Tetracentron sinense*	AM396504	Tet_sinGnBk
early diverging eudicots	Menispermaceae	*Tinospora sinensis*	EF143855	Tin_sinGnBk
early diverging eudicots	Trochodendraceae	*Trochodendron aralioides*	U92848	Tro_araGnBk
early diverging eudicots	Ranunculaceae	*Xanthorhiza simplicissima*	AF542567	Xan_simGnBk
core eudicots	Acanthaceae	*Acanthus longifolius*	AJ429326	Aca_lonGnBk
core eudicots	Aextoxicaceae	*Aextoxicon punctatum*	DQ182342	Aex_punGnBk
core eudicots	Huaceae	*Afrostyrax sp.*	EF135501	Afr_sp_GnBk
core eudicots	Cornaceae	*Alangium kurzii*	DQ341347	Ala_kurGnBk
core eudicots	Betulaceae	*Alnus japonica*	AB038176	Aln_japGnBk
core eudicots	Altingiaceae	*Altingia excelsa*	AF304520	Alt_excGnBk
core eudicots	Altingiaceae	*Altingia obovata*	AF304523	Alt_oboGnBk
core eudicots	Pseudanthaceae	*Androstachys johnsonii*	EF135502	And_johGnBk
core eudicots	Plantaginaceae	*Antirrhinum majus*	AJ429342	Ant_majGnBk
core eudicots	Aralidiaceae	*Aralidium pinnatifidum*	U58627	Ara_pinGnBk
core eudicots	Fabaceae	*Bauhinia galpinii*	EU361875	Bau_galGnBk
core eudicots	Berberidopsidaceae	*Berberidopsis corallina*	EU002171	Ber_corGnBk
core eudicots	Betulaceae	*Betula pubescens*	AY372025	Bet_pubGnBk
core eudicots	Bixaceae	*Bixa orellana*	FM179929	Bix_oreGnBk
core eudicots	Brassicaceae	*Brassica carinata*	AB354275	Bra_carGnBk
core eudicots	Bruniaceae	*Brunia albiflora*	AY490953	Bru_albGnBk
core eudicots	Rhizophoraceae	*Bruguiera gymnorhiza*	AF105088	Bru_gymGnBk
core eudicots	Theaceae	*Camellia japonica*	AF380074	Cam_japGnBk
core eudicots	Cardiopteridaceae	*Cardiopteris quinqueloba*	AJ429310	Car_quiGnBk
core eudicots	Fagaceae	*Castanea crenata*	AB107636	Cas_creGnBk

Group	Family	Species	GenBank accession ID	Taxon code
core eudicots	Icacinaceae	*Cassinopsis ilicifolia*	AJ429312	Cas_iliGnBk
core eudicots	Cannabaceae	*Celtis occidentalis*	AY257535	Cel_occGnBk
core eudicots	Cercidiphyllaceae	*Cercidiphyllum japonicum*	AM396508	Cer_japGnBk
core eudicots	Iteaceae	*Choristylis rhamnoides*	AF274609	Cho_rhaGnBk
core eudicots	Rutaceae	*Citrus sinensis*	NC_008334	Cit_sinGnBk
core eudicots	Corynocarpaceae	*Corynocarpus laevigata*	AY968448	Cor_laeGnBk
core eudicots	Hamamelidaceae	*Corylopsis sinensis*	AF013038	Cor_sinGnBk
core eudicots	Cornaceae	*Cornus walteri*	DQ340478	Cor_walGnBk
core eudicots	Crassulaceae	*Crassula orbicularis*	AF115601	Cra_orbGnBk
core eudicots	Crossosomataceae	*Crossosoma bigelovii*	DQ443456	Cro_bigGnBk
core eudicots	Cucurbitaceae	*Cucurbita pepo*	DQ536666	Cuc_pepGnBk
core eudicots	Daphniphyllaceae	*Daphniphyllum sp.*	AF274612	Dap_sp_GnBk
core eudicots	Viscaceae	*Dendrophthora clavata*	EF584636	Den_claGnBk
core eudicots	Dilleniaceae	*Dillenia indica*	DQ401359	Dil_indGnBk
core eudicots	Ebenaceae	*Diospyros texana*	DQ924060	Dio_texGnBk
core eudicots	Crassulaceae	*Dudleya viscida*	AF274614	Dud_visGnBk
core eudicots	Elaeocarpaceae	*Elaeocarpus reticulatus*	AY935931	Ela_retGnBk
core eudicots	Ericaceae	*Erica discolor*	AM889710	Eri_disGnBk
core eudicots	Geraniaceae	*Erodium cicutarium*	AM396500	Ero_cicGnBk
core eudicots	Santalaceae	*Eubrachion ambiguum*	EF464498	Eub_ambGnBk
core eudicots	Eucommiaceae	*Eucommia ulmoides*	AF345323	Euc_ulmGnBk
core eudicots	Ternstroemiaceae	*Eurya japonica*	AF380081	Eur_japGnBk
core eudicots	Hamamelidaceae	*Exbucklandia tonkinensis*	AF128832	Exb_tonGnBk
core eudicots	Fagaceae	*Fagus lucida*	EF057139	Fag_lucGnBk
core eudicots	Salicaceae	*Flacourtia jangomas*	EF135541	Fla_janGnBk
core eudicots	Fouquieriaceae	*Fouquieria diguetii*	AJ429285	Fou_digGnBk
core eudicots	Grubbiaceae	*Grubbia tomentosa*	AF323184	Gru_tomGnBk
core eudicots	Zygophyllaceae	*Guaiacum officinale*	DQ401366	Gua_offGnBk
core eudicots	Gunneraceae	*Gunnera manicata*	EU002179	Gun_manGnBk
core eudicots	Gunneraceae	*Gunnera tinctoria*	AM396506	Gun_tinGnBk
core eudicots	Haloragaceae	*Haloragis dura*	EF179019	Hal_durGnBk
core eudicots	Asteraceae	*Helianthus annuus*	NC_007977	Hel_annGnBk
core eudicots	Heteropyxidaceae	*Heteropyxis natalensis*	AF368208	Het_natGnBk
core eudicots	Saxifragaceae	*Heuchera sanguinea*	EU002180	Heu_sanGnBk
core eudicots	Celastraceae	*Hippocratea barbata*	AJ581399	Hip_barGnBk
core eudicots	Hydrangeaceae	*Hydrangea macrophylla*	AB038178	Hyd_macGnBk
core eudicots	Icacinaceae	*Icacina senegalensis*	AJ429313	Ica_senGnBk
core eudicots	Balsaminaceae	*Impatiens noli-tangere*	AF542608	Imp_notGnBk
core eudicots	Convolvulaceae	*Ipomoea batatas*	AJ429355	Ipo_batGnBk
core eudicots	Iteaceae	*Itea virginica*	EF456732	Ite_virGnBk
core eudicots	Ixerbaceae	*Ixerba brexioides*	EU002181	Ixe_breGnBk
core eudicots	Ixonanthaceae	*Ixonanthes sp.*	AB048381	Ixo_sp_GnBk
core eudicots	Juglandaceae	*Juglans nigra*	U92851	Jug_nigGnBk
core eudicots	Crassulaceae	*Kalanchoe scapigera*	AF115620	Kal_scaGnBk
core eudicots	Vitaceae	*Leea coccinea*	AM396497	Lee_cocGnBk
core eudicots	Loasaceae	*Loasa heterophylla*	AY254066	Loa_hetGnBk
core eudicots	Rubiaceae	*Luculia gratissima*	Z70199	Luc_graGnBk
core eudicots	Combretaceae	*Lumnitzera racemosa*	AB114598	Lum_racGnBk
core eudicots	Lythraceae	*Lythrum flagellare*	EU002183	Lyt_flaGnBk
core eudicots	Malpighiaceae	*Malpighia emarginata*	AF344561	Mal_emaGnBk
core eudicots	Sapotaceae	*Manilkara zapota*	DQ924092	Man_zapGnBk

Group	Family	Species	GenBank accession ID	Taxon code
core eudicots	Melianthaceae	*Melianthus comosus*	EU002184	Mel_comGnBk
core eudicots	Montiniaceae	*Montinia caryophyllacea*	AJ429359	Mnt_carGnBk
core eudicots	Myrothamnaceae	*Myrothamnus moschata*	AF542591	Myr_mosGnBk
core eudicots	Haloragaceae	*Myriophyllum sibiricum*	EF178980	Myr_sibGnBk
core eudicots	Ochnaceae	*Ochna multiflora*	EF135572	Och_mulGnBk
core eudicots	Oncothecaceae	*Oncotheca balansae*	AJ581439	Onc_balGnBk
core eudicots	Opiliacaeae	*Opilia sp.*	AY042621	Opi_sp_GnBk
core eudicots	Orobanchaceae	*Orobanche hederae*	AJ429338	Oro_hedGnBk
core eudicots	Santalaceae	*Osyris alba*	AM396499	Osy_albGnBk
core eudicots	Oxalidaceae	*Oxalis latifolia*	EU002186	Oxa_latGnBk
core eudicots	Paeoniaceae	*Paeonia suffruticosa*	AF033593	Pae_sufGnBk
core eudicots	Paracryphiaceae	*Paracryphia alticola*	AJ429367	Par_altGnBk
core eudicots	Passifloraceae	*Passiflora biflora*	EU017067	Pas_bifGnBk
core eudicots	Gesneriaceae	*Peltanthera floribunda*	AJ429330	Pel_floGnBk
core eudicots	Geraniaceae	*Pelargonium x hortorum*	NC_008454	Pel_horGnBk
core eudicots	Saxifragaceae	*Peltoboykinia tellimoides*	AB161144	Pel_telGnBk
core eudicots	Pentaphylacaceae	*Pentaphylax euryoides*	AJ429291	Pen_eurGnBk
core eudicots	Penthoraceae	*Penthorum sedoides*	EF179063	Pen_sedGnBk
core eudicots	Fabaceae	*Pisum sativum*	EU307313	Pis_satGnBk
core eudicots	Pittosporaceae	*Pittosporum undulatum*	AJ429374	Pit_undGnBk
core eudicots	Plocospermataceae	*Plocosperma buxifolium*	Z70192	Plo_buxGnBk
core eudicots	Polygalaceae	*Polygala californica*	AY386842	Pol_calGnBk
core eudicots	Primulaceae	*Primula mollis*	DQ378418	Pri_molGnBk
core eudicots	Parnassiaceae	*Parnassia grandifolia*	EF135575	Prs_graGnBk
core eudicots	Pterostemonaceae	*Pterostemon rotundifolius*	AF274630	Pte_rotGnBk
core eudicots	Rhamnaceae	*Rhamnus cathartica*	AY257533	Rha_catGnBk
core eudicots	Grossulariaceae	*Ribes aureum*	L34153	Rib_aurGnBk
core eudicots	Rousseaceae	*Roussea simplex*	AJ429389	Rou_simGnBk
core eudicots	Lamiaceae	*Salvia coccinea*	AY840147	Sal_cocGnBk
core eudicots	Lamiaceae	*Salvia splendens*	AF477765	Sal_splGnBk
core eudicots	Santalaceae	*Santalum album*	AY042650	San_albGnBk
core eudicots	Sarraceniaceae	*Sarracenia purpurea*	AJ429296	Sar_purGnBk
core eudicots	Saxifragaceae	*Saxifraga mertensiana*	L34142	Sax_merGnBk
core eudicots	Saxifragaceae	*Saxifraga stellaris*	AF115493	Sax_steGnBk
core eudicots	Schoepfiaceae	*Schoepfia schreberi*	DQ787447	Sch_schGnBk
core eudicots	Anacardiaceae	*Schinus sp.*	AY491645	Sch_sp_GnBk
core eudicots	Scrophulariaceae	*Scrophularia arguta*	AJ429349	Scr_argGnBk
core eudicots	Crassulaceae	*Sedum aizoon*	AB038187	Sed_aizGnBk
core eudicots	Sladeniaceae	*Sladenia celastrifolia*	AJ429297	Sla_celGnBk
core eudicots	Lythraceae	*Sonneratia alba*	EF408677	Son_albGnBk
core eudicots	Rosaceae	*Spiraea cantoniensis*	AF288127	Spi_canGnBk
core eudicots	Stachyuraceae	*Stachyurus praecox*	DQ443457	Sta_praGnBk
core eudicots	Styracaceae	*Styrax officinalis*	AJ429300	Sty_offGnBk
core eudicots	Saxifragaceae	*Sullivantia sullivantii*	L20130	Sul_sulGnBk
core eudicots	Symplocaceae	*Symplocos paniculata*	AY336340	Sym_panGnBk
core eudicots	Symplocaceae	*Symplocos stellaris*	AY336376	Sym_steGnBk
core eudicots	Tapisciaceae	*Tapiscia sinensis*	EU002190	Tap_sinGnBk
core eudicots	Ternstroemiaceae	*Ternstroemia longipes*	AF380110	Ter_lonGnBk
core eudicots	Dilleniaceae	*Tetracera asiatica*	AY042665	Tet_asiGnBk
core eudicots	Ebenaceae	*Tetraclis baroni*	DQ924088	Tet_barGnBk
core eudicots	Tetracarpaeaceae	*Tetracarpaea tasmanica*	L34154	Tet_tasGnBk

Group	Family	Species	GenBank accession ID	Taxon code
core eudicots	Malvaceae	*Tilia americana*	AY321191	Til_ameGnBk
core eudicots	Malvaceae	*Tilia kiusiana*	AB006386	Til_kiuGnBk
core eudicots	Tribelaceae	*Tribeles australis*	AJ429369	Tri_ausGnBk
core eudicots	Tropaeolaceae	*Tropaeolum majus*	AY483224	Trp_majGnBk
core eudicots	Vahliaceae	*Vahlia capensis*	AJ429316	Vah_capGnBk
core eudicots	Adoxaceae	*Viburnum rhytidophyllum*	AJ429391	Vib_rhyGnBk
core eudicots	Viscaceae	*Viscum articulatum*	EF464496	Vis_artGnBk
core eudicots	Vitaceae	*Vitis riparia*	AF542593	Vit_ripGnBk
core eudicots	Ulmaceae	*Zelkova schneideriana*	AF345328	Zel_schGnBk
Caryophyllales	Nyctaginaceae	*Acleisanthes somalensis*	AY042655	Acl_somGnBk
Caryophyllales	Molluginaceae	*Adenogramma sp.*	AY042535	Ade_sp_GnBk
Caryophyllales	Phytolaccaceae ll	*Agdestis clematidea*	AY042538	Agd_cleGnBk
Caryophyllales	Caryophyllaceae	*Agrostemma githago*	AY042539	Agr_gitGnBk
Caryophyllales	Didiereaceae	*Alluaudiopsis fiherenensis*	AY042542	Ald_fihGnBk
Caryophyllales	Didiereaceae	*Alluaudia ascendens*	AY042541	All_ascGnBk
Caryophyllales	Nyctaginaceae	*Allionia incarnata*	AY042540	All_incGnBk
Caryophyllales	Amaranthaceae	*Amaranthus greggii*	AY514808	Ama_greGnBk
Caryophyllales	Anacampserotaceae	*Anacampseros karasmontana*	DQ855859	Ana_karGnBk
Caryophyllales	Anacampserotaceae	*Anacampseros retusa*	DQ855860	Ana_retGnBk
Caryophyllales	Anacampserotaceae	*Anacampseros subnuda*	DQ855861	Ana_subGnBk
Caryophyllales	Anacampserotaceae	*Anacampseros telephiastrum*	DQ855862	Ana_telGnBk
Caryophyllales	Ancistrocladaceae	*Ancistrocladus abbreviatus*	AF315939	Anc_abbGnBk
Caryophyllales	Ancistrocladaceae	*Ancistrocladus hamatus*	AF204842	Anc_hamGnBk
Caryophyllales	Ancistrocladaceae	*Ancistrocladus heyneanus*	AF204841	Anc_heyGnBk
Caryophyllales	Ancistrocladaceae	*Ancistrocladus korupensis*	AY042546	Anc_korGnBk
Caryophyllales	Basellaceae	*Anredera cordifolia*	AY042547	Anr_corGnBk
Caryophyllales	Caryophyllaceae	*Arenaria koriniana*	AY936318	Are_korGnBk
Caryophyllales	Caryophyllaceae	*Arenaria nevadensis*	AY936303	Are_nevGnBk
Caryophyllales	Aizoaceae	*Aridaria noctiflora*	AY042619	Ari_nocGnBk
Caryophyllales	Asteropeiaceae	*Asteropeia micraster*	AY042549	Ast_micGnBk
Caryophyllales	Amaranthaceae	*Atriplex patula*	AY042550	Atr_patGnBk
Caryophyllales	Amaranthaceae	*Atriplex tornabeni*	AY936329	Atr_torGnBk
Caryophyllales	Anacampserotaceae	*Avonia albissima*	DQ855856	Avo_albGnBk
Caryophyllales	Anacampserotaceae	*Avonia papyracea*	AY042545	Avo_papGnBk
Caryophyllales	Anacampserotaceae	*Avonia recurvata*	DQ855858	Avo_recGnBk
Caryophyllales	Barbeuiaceae	*Barbeuia madagascariensis*	AY042552	Bar_madGnBk
Caryophyllales	Basellaceae	*Basella alba*	AY042553	Bas_albGnBk
Caryophyllales	Nyctaginaceae	*Boerhavia coccinea*	AY042558	Boe_cocGnBk
Caryophyllales	Nyctaginaceae	*Bougainvillea glabra*	AY042560	Bou_glaGnBk
Caryophyllales	Nyctaginaceae	*Bougainvillea sp.*	AF204865	Bou_sp_GnBk
Caryophyllales	Montiaceae	*Calandrinia ciliata*	AY764127	Cal_cilGnBk
Caryophyllales	Montiaceae	*Calandrinia feltonii*	AY042562	Cal_felGnBk
Caryophyllales	Didiereaceae	*Calyptrotheca somalensis*	AY042563	Cal_somGnBk
Caryophyllales	Amaranthaceae	*Celosia sp.*	AY042565	Cel_sp_GnBk
Caryophyllales	Amaranthaceae	*Celosia trigyna*	AY514811	Cel_triGnBk
Caryophyllales	Didiereaceae	*Ceraria fruticulosa*	AY875371	Cer_fruGnBk
Caryophyllales	Amaranthaceae	*Chenopodium acuminatum*	AY514836	Che_acuGnBk
Caryophyllales	Amaranthaceae	*Chenopodium bonus-henricus*	AY514834	Che_bohGnBk
Caryophyllales	Amaranthaceae	*Chenopodium botrys*	AY514835	Che_botGnBk
Caryophyllales	Montiaceae	*Cistanthe grandiflora*	AY042568	Cis_graGnBk
Caryophyllales	Montiaceae	*Claytonia acutifolia*	AY764097	Cly_acuGnBk

Group	Family	Species	GenBank accession ID	Taxon code
Caryophyllales	Montiaceae	*Claytonia arenicola*	AY764088	Cly_areGnBk
Caryophyllales	Montiaceae	*Claytonia gypsophiloides*	AY764090	Cly_gypGnBk
Caryophyllales	Montiaceae	*Claytonia megarhiza*	AY042569	Cly_megGnBk
Caryophyllales	Montiaceae	*Claytonia nevadensis*	AY764104	Cly_nevGnBk
Caryophyllales	Montiaceae	*Claytonia ogilviensis*	AY764105	Cly_ogiGnBk
Caryophyllales	Montiaceae	*Claytonia perfoliata*	AY764091	Cly_perGnBk
Caryophyllales	Montiaceae	*Claytonia sarmentosa*	AY764107	Cly_sarGnBk
Caryophyllales	Montiaceae	*Claytonia saxosa*	AY764094	Cly_saxGnBk
Caryophyllales	Montiaceae	*Claytonia scammaniana*	AY764108	Cly_scaGnBk
Caryophyllales	Montiaceae	*Claytonia sibirica*	AY764109	Cly_sibGnBk
Caryophyllales	Montiaceae	*Claytonia tuberosa*	AY764110	Cly_tubGnBk
Caryophyllales	Montiaceae	*Claytonia umbellata*	AY764112	Cly_umbGnBk
Caryophyllales	Montiaceae	*Claytonia virginica*	AY764113	Cly_virGnBk
Caryophyllales	Montiaceae	*Claytonia washingtoniana*	AY764095	Cly_wasGnBk
Caryophyllales	Polygonaceae	*Coccoloba pyrifolia*	EF437994	Coc_pyrGnBk
Caryophyllales	Polygonaceae	*Coccoloba swartzii*	EF437995	Coc_swaGnBk
Caryophyllales	Polygonaceae	*Coccoloba uvifera*	EF437996	Coc_uviGnBk
Caryophyllales	Nyctaginaceae	*Commicarpus raynalii*	AY042571	Com_rayGnBk
Caryophyllales	Molluginaceae	*Corbichonia decumbens*	AY042572	Cor_decGnBk
Caryophyllales	Didiereaceae	*Decarya madagascariensis*	AY042574	Dec_madGnBk
Caryophyllales	Aizoaceae	*Delosperma cooperi*	DQ855843	Del_cooGnBk
Caryophyllales	Aizoaceae	*Delosperma echinatum*	AY042575	Del_echGnBk
Caryophyllales	Caryophyllaceae	*Dianthus seguieri*	AY936321	Dia_segGnBk
Caryophyllales	Didiereaceae	*Didierea trollii*	AY042576	Did_troGnBk
Caryophyllales	Droseraceae	*Drosera adelae*	AY096121	Dro_adeGnBk
Caryophyllales	Droseraceae	*Drosera aliciae*	AF204849	Dro_aliGnBk
Caryophyllales	Droseraceae	*Drosera capensis*	AY096122	Dro_capGnBk
Caryophyllales	Droseraceae	*Drosera communis*	AY042579	Dro_comGnBk
Caryophyllales	Droseraceae	*Drosera capillaris*	AF204850	Dro_cplGnBk
Caryophyllales	Droseraceae	*Drosophyllum lusitanicum*	AY514860	Dro_lusGnBk
Caryophyllales	Droseraceae	*Drosera regia*	AF204848	Dro_regGnBk
Caryophyllales	Caryophyllaceae	*Drypis spinosa*	AY936293	Dry_spiGnBk
Caryophyllales	Phytolaccaceae I	*Ercilla volubilis*	AY042583	Erc_volGnBk
Caryophyllales	Frankeniaceae	*Frankenia corymbosa*	AY042587	Frk_corGnBk
Caryophyllales	Frankeniaceae	*Frankenia laevis*	AY514853	Frk_laeGnBk
Caryophyllales	Amaranthaceae	*Froelichia floridana*	AY514799	Fro_floGnBk
Caryophyllales	Amaranthaceae	*Froelichia gracilis*	AY042588	Fro_graGnBk
Caryophyllales	Phytolaccaceae II	*Gallesia integrifolia*	AY042590	Gal_intGnBk
Caryophyllales	Aizoaceae	*Galenia pubescens*	AY042589	Gal_pubGnBk
Caryophyllales	Gisekiaceae	*Gisekia africana*	AY042591	Gis_afrGnBk
Caryophyllales	Molluginaceae	*Glinus lotoides*	AY042592	Gli_lotGnBk
Caryophyllales	Molluginaceae	*Glischrothamnus ulei*	AY042593	Gls_uleGnBk
Caryophyllales	Anacampserotaceae	*Grahamia australiana*	DQ855855	Gra_ausGnBk
Caryophyllales	Anacampserotaceae	*Grahamia bracteata*	AY015273	Gra_braGnBk
Caryophyllales	Anacampserotaceae	*Grahamia coahuilensis*	DQ855854	Gra_coaGnBk
Caryophyllales	Anacampserotaceae	*Grahamia frutescens*	DQ855851	Gra_fruGnBk
Caryophyllales	Anacampserotaceae	*Grahamia kurtzii*	DQ855853	Gra_kurGnBk
Caryophyllales	Anacampserotaceae	*Grahamia vulcanensis*	DQ855852	Gra_vulGnBk
Caryophyllales	Halophytaceae	*Halophytum ameghinoi*	AY514852	Hal_ameGnBk
Caryophyllales	Montiaceae	*Hectorella caespitosa*	EF551350	Hec_caeGnBk
Caryophyllales	Caryophyllaceae	*Herniaria baetica*	AY936283	Her_baeGnBk

Group	Family	Species	GenBank accession ID	Taxon code
Caryophyllales	Phytolaccaceae II	*Hilleria latifolia*	AY042601	Hil_latGnBk
Caryophyllales	Phytolaccaceae II	*Ledenbergia macrantha*	AY042606	Led_macGnBk
Caryophyllales	Montiaceae	*Lewisia cantelovii*	AY042607	Lew_canGnBk
Caryophyllales	Montiaceae	*Lewisia columbiana*	AY764126	Lew_colGnBk
Caryophyllales	Montiaceae	*Lewisia rediviva*	AY764125	Lew_redGnBk
Caryophyllales	Molluginaceae	*Limeum africanum*	AY042608	Lim_afrGnBk
Caryophyllales	Plumbaginaceae	*Limonium cavanillesii*	AY042610	Lmn_cavGnBk
Caryophyllales	Plumbaginaceae	*Limonium latifolium*	AY514861	Lmn_latGnBk
Caryophyllales	Plumbaginaceae	*Limonium rigualii*	AM889717	Lmn_rigGnBk
Caryophyllales	Plumbaginaceae	*Limonium thiniense*	AM889718	Lmn_thiGnBk
Caryophyllales	Lophiocarpaceae	*Lophiocarpus burchellii*	AY042611	Lop_burGnBk
Caryophyllales	Montiaceae	*Lyallia kerguelensis*	EF551349	Lya_kerGnBk
Caryophyllales	Caryophyllaceae	*Minuartia geniculata*	AY936307	Min_genGnBk
Caryophyllales	Nyctaginaceae	*Mirabilis jalapa*	AY042614	Mir_jalGnBk
Caryophyllales	Nyctaginaceae	*Mirabilis nyctaginea*	AY042624	Mir_nycGnBk
Caryophyllales	Caryophyllaceae	*Moehringia intricata*	AY936305	Moe_intGnBk
Caryophyllales	Caryophyllaceae	*Moehringia trinervia*	AY042615	Moe_triGnBk
Caryophyllales	Molluginaceae	*Mollugo verticillata*	AY936330	Mol_verGnBk
Caryophyllales	Montiaceae	*Montia bostockii*	AY764114	Mon_bosGnBk
Caryophyllales	Montiaceae	*Montia diffusa*	AY764121	Mon_difGnBk
Caryophyllales	Montiaceae	*Montia fontana*	AY764119	Mon_fonGnBk
Caryophyllales	Montiaceae	*Montia howellii*	AY764117	Mon_howGnBk
Caryophyllales	Montiaceae	*Montia linearis*	AY764116	Mon_linGnBk
Caryophyllales	Montiaceae	*Montia parvifolia*	AY764122+	Mon_parGnBk
Caryophyllales	Nepenthaceae	*Nepenthes alata*	DQ991360	Nep_alaGnBk
Caryophyllales	Nepenthaceae	*Nepenthes albomarginata*	DQ991358	Nep_albGnBk
Caryophyllales	Nepenthaceae	*Nepenthes burkei*	DQ840247	Nep_burGnBk
Caryophyllales	Nepenthaceae	*Nepenthes edwardsiana*	DQ840248	Nep_edwGnBk
Caryophyllales	Nepenthaceae	*Nepenthes sibuyanensis*	DQ840246	Nep_sibGnBk
Caryophyllales	Montiaceae	*Neopaxia erythrophylla*	AY764123	Npx_eryGnBk
Caryophyllales	Montiaceae	*Neopaxia racemosa*	AY764124	Npx_racGnBk
Caryophyllales	Caryophyllaceae	*Ortegia hispanica*	AY936286	Ort_hisGnBk
Caryophyllales	Caryophyllaceae	*Paronychia echinulata*	AY936285	Par_echGnBk
Caryophyllales	Phytolaccaceae II	*Petiveria alliacea*	AY042628	Pet_allGnBk
Caryophyllales	Molluginaceae	*Pharnaceum sp.*	AY042629	Pha_sp_GnBk
Caryophyllales	Achatocarpaceae	*Phaulothamnus spinescens*	AY514846	Pha_spiGnBk
Caryophyllales	Phytolaccaceae I	*Phytolacca americana*	DQ401362	Phy_ameGnBk
Caryophyllales	Phytolaccaceae I	*Phytolacca dioica*	AY042631	Phy_dioGnBk
Caryophyllales	Nyctaginaceae	*Pisonia umbellifera*	AY042632	Pis_umbGnBk
Caryophyllales	Aizoaceae	*Plinthus cryptocarpus*	AY042633	Pli_cryGnBk
Caryophyllales	Plumbaginaceae	*Plumbago auriculata*	EU002187	Plu_aurGnBk
Caryophyllales	Plumbaginaceae	*Plumbago europaea*	AY042634	Plu_eurGnBk
Caryophyllales	Plumbaginaceae	*Plumbago indica*	AF204857	Plu_indGnBk
Caryophyllales	Polygonaceae	*Polygonum amphibium*	EF653725	Pol_ampGnBk
Caryophyllales	Polygonaceae	*Polygonum aviculare*	EF438020	Pol_aviGnBk
Caryophyllales	Polygonaceae	*Polygonum filiforme*	EF653722	Pol_filGnBk
Caryophyllales	Polygonaceae	*Polygonum forrestii*	EF438012	Pol_forGnBk
Caryophyllales	Polygonaceae	*Polygonum paraguayense*	EU196952	Pol_parGnBk
Caryophyllales	Caryophyllaceae	*Polycarpon tetraphyllum*	AY936287	Pol_tetGnBk
Caryophyllales	Portulacaceae	*Portulaca bicolor*	DQ855848	Por_bicGnBk
Caryophyllales	Portulacaceae	*Portulaca eruca*	DQ855849	Por_eruGnBk

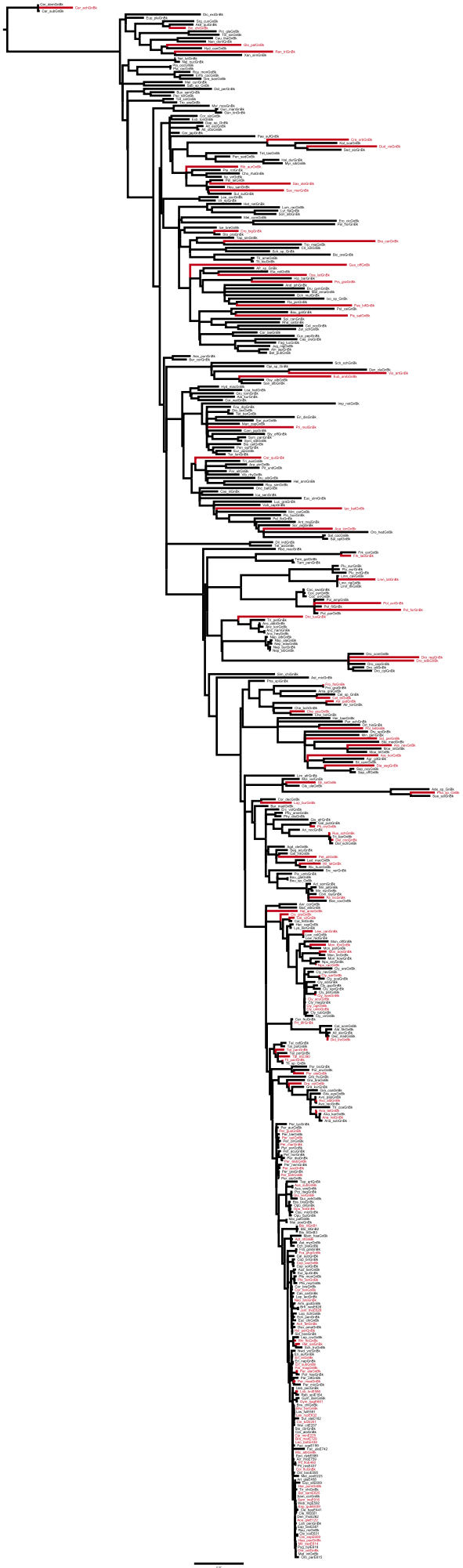
Group	Family	Species	GenBank accession ID	Taxon code
Caryophyllales	Portulacaceae	*Portulaca oleracea*	AY875349	Por_oleGnBk
Caryophyllales	Didiereaceae	*Portulacaria afra*	AY875368	Prt_afrGnBk
Caryophyllales	Rhabdodendraceae	*Rhabdodendron macrophyllum*	AY042642	Rbd_macGnBk
Caryophyllales	Phytolaccaceae ll	*Rivina humilis*	AY514850	Riv_humGnBk
Caryophyllales	Aizoaceae	*Ruschia schollii*	AY042649	Rus_schGnBk
Caryophyllales	Caryophyllaceae	*Saponaria ocymoides*	AY042651	Sap_ocyGnBk
Caryophyllales	Caryophyllaceae	*Saponaria officinalis*	AY936325	Sap_offGnBk
Caryophyllales	Caryophyllaceae	*Scleranthus perennis*	AY514847	Scl_perGnBk
Caryophyllales	Phytolaccaceae ll	*Seguieria aculeata*	AY042654	Seg_acuGnBk
Caryophyllales	Caryophyllaceae	*Silene campanula*	AY936311	Sil_camGnBk
Caryophyllales	Simmondsiaceae	*Simmondsia chinensis*	AY042657	Sim_chiGnBk
Caryophyllales	Sarcobataceae	*Sarcobatus vermiculatus*	AY042652	Src_verGnBk
Caryophyllales	Caryophyllaceae	*Stellaria media*	AY936299	Ste_medGnBk
Caryophyllales	Molluginaceae	*Suessenguthiella scleranthoides*	AY042659	Sue_sclGnBk
Caryophyllales	Talinaceae	*Talinum caffrum*	AY042662	Tal_cafGnBk
Caryophyllales	Talinaceae	*Talinum paniculatum*	AY015274	Tal_panGnBk
Caryophyllales	Talinaceae	*Talinum polygaloides*	DQ855845	Tal_polGnBk
Caryophyllales	Talinaceae	*Talinum portulacifolium*	DQ855847	Tal_porGnBk
Caryophyllales	Talinaceae	*Talinum triangulare*	-	Tal_triE390
Caryophyllales	Tamaricaceae	*Tamarix gallica*	AF204861	Tam_galGnBk
Caryophyllales	Tamaricaceae	*Tamarix pentandra*	AY042663	Tam_penGnBk
Caryophyllales	Talinaceae	*Talinella pachypoda*	DQ855846	Tll_pacGnBk
Caryophyllales	Talinaceae	*Talinella sp.*	AY514859	Tll_sp_GnBk
Caryophyllales	Anacampserotaceae	*Talinaria coahuilensis*	AY042661	Tlr_coaGnBk
Caryophyllales	Aizoaceae	*Trichodiadema barbatum*	AY042666	Tri_barGnBk
Caryophyllales	Dioncophyllaceae	*Triphyophyllum peltatum*	AF315940	Tri_pelGnBk
Cactaceae	Cactaceae	*Acanthocalycium thionanthum*	-	Aca_glaE122
Cactaceae	Cactaceae	*Acanthocereus pentagonus*	AY015295	Ach_tetGnBk
Cactaceae	Cactaceae	*Armatocereus godingianus*	AY015296	Arm_godGnBk
Cactaceae	Cactaceae	*Arrojadoa rhodantha*	JX683842	Arr_rhoE759
Cactaceae	Cactaceae	*Arthrocereus glaziovii*	JX683846	Art_glaE458
Cactaceae	Cactaceae	*Astrophytum myriostigma*	AY015288	Ast_myrGnBk
Cactaceae	Cactaceae	*Austrocactus bertinii*	AY015300	Aus_berGnBk
Cactaceae	Cactaceae	*Austrocylindropuntia subulata*	AY875364	Aus_subGnBk
Cactaceae	Cactaceae	*Austrocylindropuntia vestita*	AY015278	Aus_vesGnBk
Cactaceae	Cactaceae	*Aztekium ritteri*	AY015290	Azt_ritGnBk
Cactaceae	Cactaceae	*Blossfeldia liliputana*	AY875366	Blo_lilGnB1
Cactaceae	Cactaceae	*Blossfeldia liliputana*	AY015284	Blo_lilGnB2
Cactaceae	Cactaceae	*Blossfeldia liliputana*	AY015283	Blo_lilGnB3
Cactaceae	Cactaceae	*Cleistocactus samaipatanus*	JX683873	Bol_samE625
Cactaceae	Cactaceae	*Brasiliopuntia brasiliensis*	AY875370	Bra_braGnBk
Cactaceae	Cactaceae	*Brasilicereus markgrafii*	JX683870	Bra_marE729
Cactaceae	Cactaceae	*Brachycereus nesioticus*	JX683859	Brh_nesE626
Cactaceae	Cactaceae	*Browningia chlorocarpa*	AY015316	Brw_chlGnBk
Cactaceae	Cactaceae	*Browningia hertlingiana*	AY015315	Brw_herGnBk
Cactaceae	Cactaceae	*Calymmanthium substerile*	AY015291	Cal_subGnBk
Cactaceae	Cactaceae	*Castellanosia caineana*	AY015298	Cas_caiGnBk
Cactaceae	Cactaceae	*Cereus alacriportanus*	AY015313	Cer_alaGnBk
Cactaceae	Cactaceae	*Cintia knizei*	JX683861	Cin_kniE201
Cactaceae	Cactaceae	*Cipocerreus minens*	JX683867	Cip_minE225
Cactaceae	Cactaceae	*Cleistocactus baumannii*	JX683877	Cle_bauE441

Group	Family	Species	GenBank accession ID	Taxon code
Cactaceae	Cactaceae	*Borzicactus icosagonus*	JX683866	Cle_icoE631
Cactaceae	Cactaceae	*Cleistocactus ritteri*	JX683864	Cle_ritE601
Cactaceae	Cactaceae	*Borzicactus sepium*	JX683852	Cle_sepE608
Cactaceae	Cactaceae	*Coleocephalocereus fluminensis*	AY015318	Col_fluGnBk
Cactaceae	Cactaceae	*Copiapoa bridgesii*	AY015293	Cop_briGnBk
Cactaceae	Cactaceae	*Copiapoa laui*	AY015294	Cop_lauGnBk
Cactaceae	Cactaceae	*Copiapoa solaris*	AY015292	Cop_solGnBk
Cactaceae	Cactaceae	*Corryocactus brevistylus*	AY015302	Cor_breGnBk
Cactaceae	Cactaceae	*Corryocactus tenuiculus*	AY015303	Cor_tenGnBk
Cactaceae	Cactaceae	*Denmoza rhodacantha*	JX683840	Den_rhoE262
Cactaceae	Cactaceae	*Discocactus zehntneri*	JX683848	Dis_booE595
Cactaceae	Cactaceae	*Disocactus amazonicus*	AY015312	Dso_amaGnBk
Cactaceae	Cactaceae	*Echinocereus pentalophus*	AY015307	Ech_penGnBk
Cactaceae	Cactaceae	*Echinocactus platyacanthus*	AY015287	Ech_plaGnBk
Cactaceae	Cactaceae	*Lobivia chamaecereus*	JX683860	Ecp_silE609
Cactaceae	Cactaceae	*Espostoopsis dybowskii*	JX683854	Epo_dybE589
Cactaceae	Cactaceae	*Eriosyce aurata*	AY015336	Eri_aurGnBk
Cactaceae	Cactaceae	*Eriosyce islayensis*	AY015337	Eri_islGnBk
Cactaceae	Cactaceae	*Eriosyce napina*	AY015339	Eri_napGnBk
Cactaceae	Cactaceae	*Eriosyce subgibbosa*	AY015338	Eri_subGnBk
Cactaceae	Cactaceae	*Escontria chiotilla*	AY015308	Esc_chiGnBk
Cactaceae	Cactaceae	*Vatricania guentheri*	JX683871	Esp_gueE639
Cactaceae	Cactaceae	*Espostoa lanata*	JX683863	Esp_lanE587
Cactaceae	Cactaceae	*Eulychnia iquiquensis*	AY015301	Eul_iquGnBk
Cactaceae	Cactaceae	*Facheiroa squamosa*	JX683865	Fac_sqaE199
Cactaceae	Cactaceae	*Facheiroa ulei*	JX683841	Fac_uleE742
Cactaceae	Cactaceae	*Frailea gracillima*	AY015285	Fra_graGnBk
Cactaceae	Cactaceae	*Frailea phaeodisca*	AY015286	Fra_phaGnBk
Cactaceae	Cactaceae	*Gymnocalycium denudatum*	AY015317	Gym_denGnBk
Cactaceae	Cactaceae	*Gymnocalycium saglione*	JX683853	Gym_sagE651
Cactaceae	Cactaceae	*Haageocereus pseudomelanostele*	AY015329	Haa_pseGnBk
Cactaceae	Cactaceae	*Harrisia pomanensis*	AY015324	Har_pomGnBk
Cactaceae	Cactaceae	*Hatiora salicornioides*	AY015341	Hat_salGnBk
Cactaceae	Cactaceae	*Hylocereus peruvianus*	AY015310	Hyl_perGnBk
Cactaceae	Cactaceae	*Jasminocereus thouarsii*	JX683856	Jsm_thoE628
Cactaceae	Cactaceae	*Lasiocereus fulvus*	JX683843	Las_fulE591
Cactaceae	Cactaceae	*Lasiocereus rupicola*	JX683844	Las_rupE632
Cactaceae	Cactaceae	*Leocereus bahiensis*	JX683874	Leo_bahE438
Cactaceae	Cactaceae	*Lepismium cruciforme*	AY015344	Lep_cruGnBk
Cactaceae	Cactaceae	*Leptocereus leonii*	AY015297	Lep_leoGnBk
Cactaceae	Cactaceae	*Mediolobivia pygmaea*	JX683851	Lob_kniE586
Cactaceae	Cactaceae	*Lobivia pentlandii*	AY015323	Lob_penGnBk
Cactaceae	Cactaceae	*Pachycereus schottii*	AY015309	Lop_schGnBk
Cactaceae	Cactaceae	*Maihuenia patagonica*	AY015281	Mai_patGnBk
Cactaceae	Cactaceae	*Maihuenia poeppigii*	AY015282	Mai_poeGnBk
Cactaceae	Cactaceae	*Mammillaria haageana*	AY015289	Mam_haaGnBk
Cactaceae	Cactaceae	*Matucana intertexta*	AY015327	Mat_intGnBk
Cactaceae	Cactaceae	*Melocactus zehntneri*	JX683849	Mel_zenE223
Cactaceae	Cactaceae	*Micranthocereus albicephalus*	AY015314	Mic_albGnBk
Cactaceae	Cactaceae	*Mila caespitosa*	JX683872	Mil_cspE614
Cactaceae	Cactaceae	*Neoraimondia arequipensis*	AY015299	Neo_areGnBk

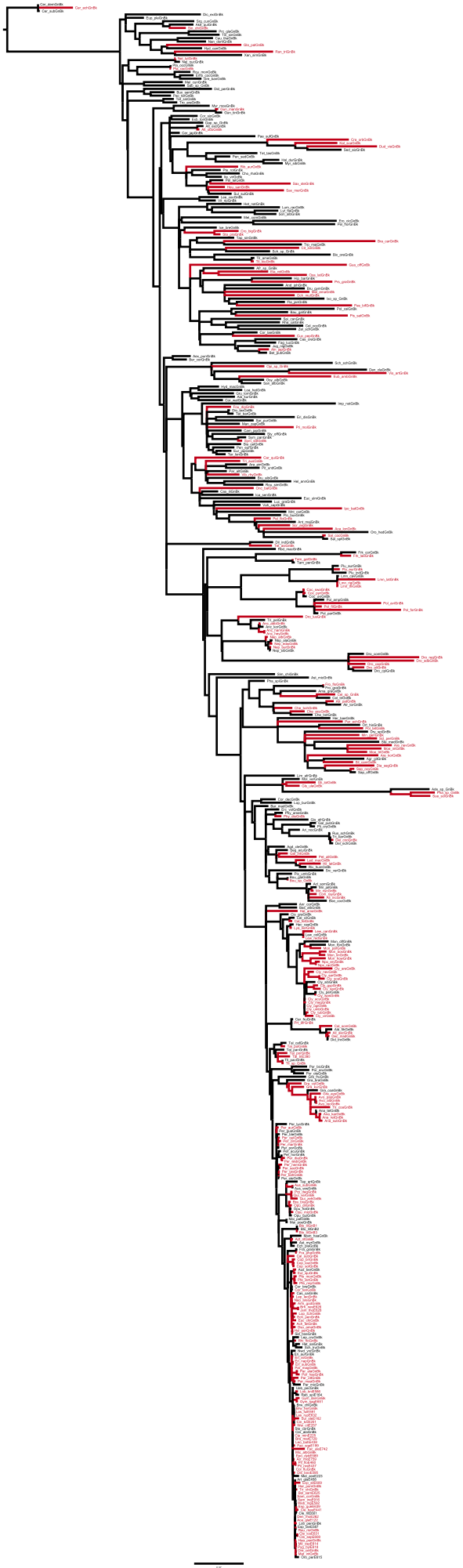
Group	Family	Species	GenBank accession ID	Taxon code
Cactaceae	Cactaceae	*Neowerdermannia vorwerkii*	AY015340	Nwd_vorGnBk
Cactaceae	Cactaceae	*Opuntia dillenii*	AY875369	Opu_dilGnBk
Cactaceae	Cactaceae	*Opuntia fragilis*	EF590413	Opu_fraGnBk
Cactaceae	Cactaceae	*Opuntia microdasys*	AY042622	Opu_micGnBk
Cactaceae	Cactaceae	*Opuntia quimilo*	AY015279	Opu_quiGnBk
Cactaceae	Cactaceae	*Oreocereus celsianus*	AY015328	Ore_celGnBk
Cactaceae	Cactaceae	*Oroya peruviana*	JX683875	Oro_perE615
Cactaceae	Cactaceae	*Parodia buenekeri*	AY015331	Par_alaGnBk
Cactaceae	Cactaceae	*Parodia haselbergii*	AY015330	Par_hasGnBk
Cactaceae	Cactaceae	*Parodia maassii*	AY015333	Par_maaGnBk
Cactaceae	Cactaceae	*Parodia magnifica*	AY015332	Par_magGnBk
Cactaceae	Cactaceae	*Parodia microsperma*	AY015334	Par_micGnBk
Cactaceae	Cactaceae	*Parodia ottonis*	AY015335	Par_ottGnBk
Cactaceae	Cactaceae	*Pereskia aculeata*	DQ855863	Per_acuGnBk
Cactaceae	Cactaceae	*Pereskia aureiflora*	AY875354	Per_aurGnBk
Cactaceae	Cactaceae	*Pereskia bahiensis*	AY875351	Per_bahGnBk
Cactaceae	Cactaceae	*Pereskia bleo*	AY875359	Per_bleGnBk
Cactaceae	Cactaceae	*Pereskia diaz-romeroana*	AY875353	Per_diaGnBk
Cactaceae	Cactaceae	*Pereskia grandifolia*	AY875362	Per_graGnBk
Cactaceae	Cactaceae	*Pereskia guamacho*	AY015275	Per_guaGnBk
Cactaceae	Cactaceae	*Pereskia horrida*	AY875356	Per_horGnBk
Cactaceae	Cactaceae	*Pereskia lychnidiflora*	AY875358	Per_lycGnBk
Cactaceae	Cactaceae	*Pereskia marcanoi*	AY875360	Per_marGnBk
Cactaceae	Cactaceae	*Pereskia nemorosa*	AY875350	Per_nemGnBk
Cactaceae	Cactaceae	*Pereskia portulacifolia*	AY875361	Per_porGnBk
Cactaceae	Cactaceae	*Pereskia quisqueyana*	AY875352	Per_quiGnBk
Cactaceae	Cactaceae	*Pereskia sacharosa*	AY875363	Per_sacGnBk
Cactaceae	Cactaceae	*Pereskia stenantha*	AY015276	Per_steGnBk
Cactaceae	Cactaceae	*Pereskia weberiana*	AY875357	Per_webGnBk
Cactaceae	Cactaceae	*Pereskia zinniiflora*	AY015277	Per_zinGnBk
Cactaceae	Cactaceae	*Pfeiffera ianthothele*	AY015304	Pfe_ianGnBk
Cactaceae	Cactaceae	*Pfeiffera miyagawae*	AY015305	Pfe_miyGnBk
Cactaceae	Cactaceae	*Pfeiffera monacantha*	AY015306	Pfe_monGnBk
Cactaceae	Cactaceae	*Pilosocereus floccosus*	JX683847	Pil_floE460
Cactaceae	Cactaceae	*Pilosocereus rosae*	JX683850	Pil_rosE437
Cactaceae	Cactaceae	*Pereskiopsis deguetii*	AY015280	Prp_degGnBk
Cactaceae	Cactaceae	*Pygmaeocereus bylesianus*	JX683862	Pyg_bylE618
Cactaceae	Cactaceae	*Quiabentia verticillata*	AY042641	Qui_verGnBk
Cactaceae	Cactaceae	*Quiabentia zehntneri*	AY875372	Qui_zehGnBk
Cactaceae	Cactaceae	*Rauhocereus riosaniensis*	AY015326	Rau_rioGnBk
Cactaceae	Cactaceae	*Aylostera fiebrigii*	JX683857	Reb_spiE164
Cactaceae	Cactaceae	*Rhipsalis floccosa*	AY015342	Rhi_floGnBk
Cactaceae	Cactaceae	*Samaipaticereus corroanus*	AY015321	Sam_corGnBk
Cactaceae	Cactaceae	*Yungasocereus inquisivensis*	JX683858	Sam_inqE916
Cactaceae	Cactaceae	*Schlumbergera truncata*	AY015343	Sch_truGnBk
Cactaceae	Cactaceae	*Selenicereus boeckmannii*	AY015311	Sel_boeGnBk
Cactaceae	Cactaceae	*Stetsonia coryne*	AY015320	Ste_corGnBk
Cactaceae	Cactaceae	*Weingartia steinbachii*	JX683868	Sul_steE162
Cactaceae	Cactaceae	*Tephrocactus articulatus*	AY875367	Tep_artGnBk
Cactaceae	Cactaceae	*Echinopsis chiloensis*	AY015322	Tri_chiGnBk
Cactaceae	Cactaceae	*Uebelmannia pectinifera*	AY015319	Ueb_pecGnBk

Group	Family	Species	GenBank accession ID	Taxon code
Cactaceae	Cactaceae	*Weberbauerocereus albus*	JX683869	Web_lngE592
Cactaceae	Cactaceae	*Weingarthia cintiensis*	JX683855	Wei_citE257

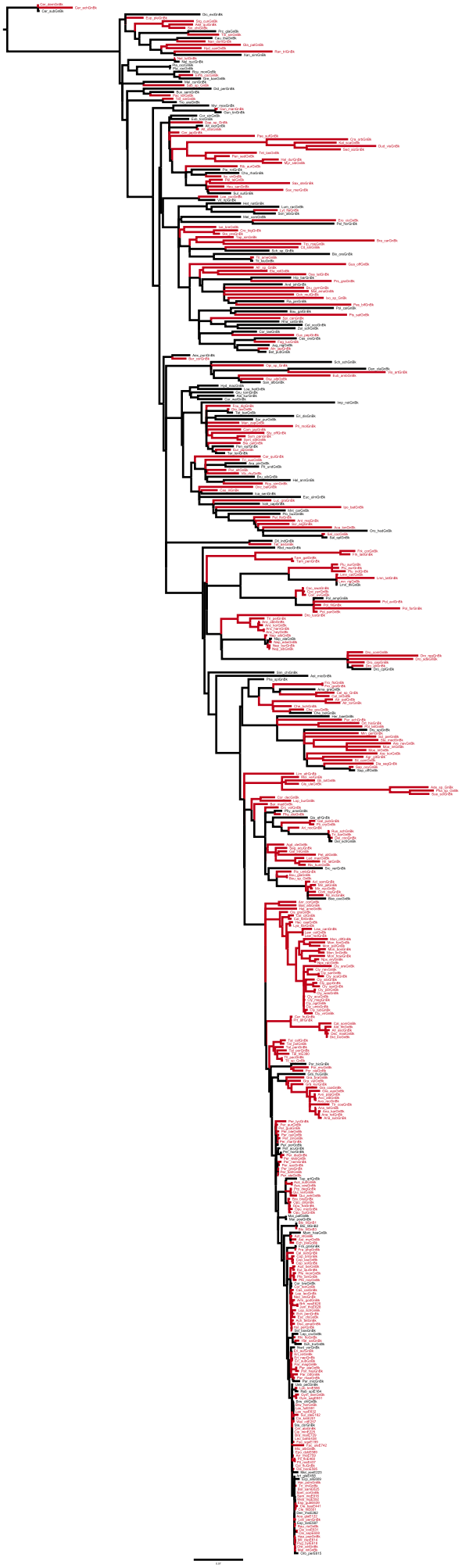
Supplement S4a: dsB



Supplement S4b: dsC



Supplement S4c: dsD

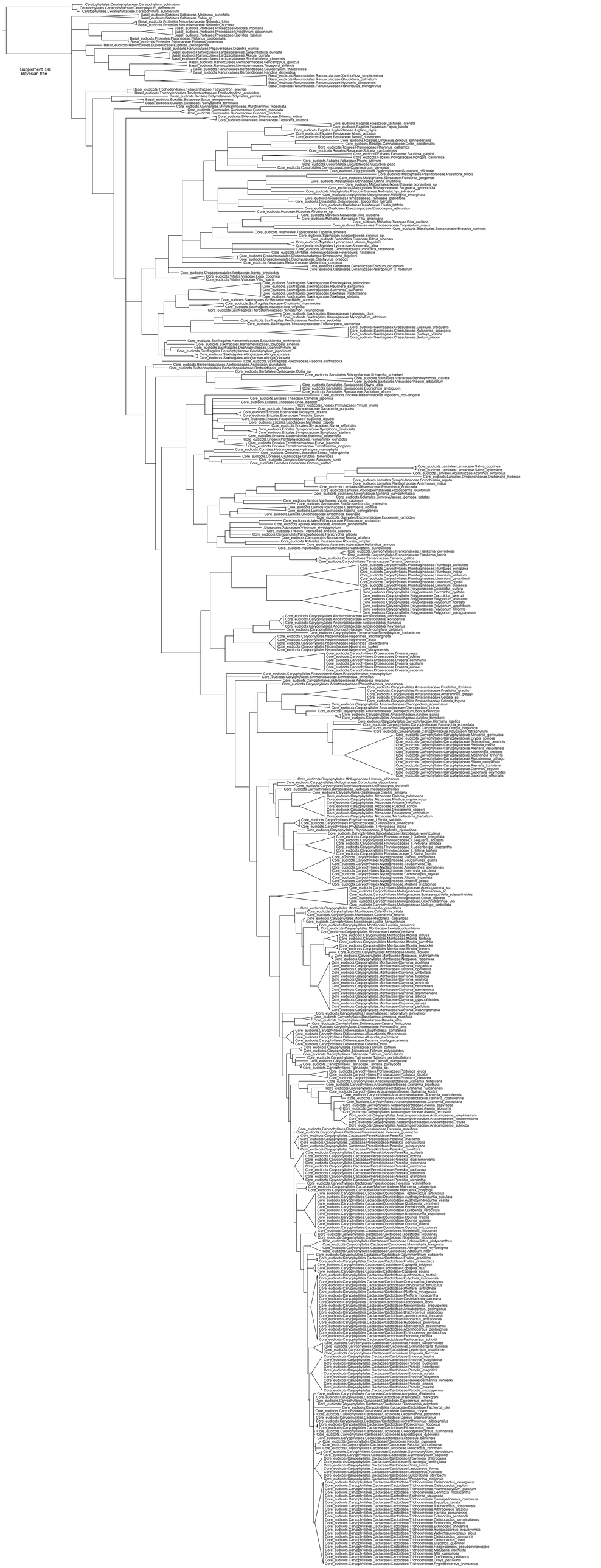


Supplement S5: Number of taxa in clades

clade	dsA	dsB	dsC	dsD	dsE	dsF	dsG	dsH	Reference for the phylogenetic placement of the clade	Reference for the phylogenetic relationships within the clade
Ceratophyllales	3	2	2	1	3	1	pruned	pruned	Soltis et al., 2007	Soltis et al., 2007
Ranunculales	13	10	10	4	13	4	pruned	pruned	Soltis et al., 2007	Soltis et al., 2007
Proteales	7	7	7	5	7	5	pruned	pruned	Soltis et al., 2007	Worberg et al., 2007; Soltis et al., 2007
Sabiales	2	2	2	1	2	1	pruned	pruned	Soltis et al., 2007	Soltis et al., 2007
Buxales	3	3	3	2	3	2	pruned	pruned	Soltis et al., 2007	Soltis et al., 2007
Trochodendrales	2	2	2	1	2	1	pruned	pruned	Soltis et al., 2007	Soltis et al., 2007
Gunnerales	3	3	3	2	3	2	pruned	pruned	Soltis et al., 2007	Soltis et al., 2007
Fagales	5	5	5	3	5	3	pruned	pruned	Soltis et al., 2007	Soltis et al., 2007
Cucurbitales	2	2	1	1	2	1	pruned	pruned	Soltis et al., 2007	Soltis et al., 2007
Rosales	4	4	4	3	4	3	pruned	pruned	Soltis et al., 2007	Soltis et al., 2007
Fabales	3	2	2	2	3	2	pruned	pruned	Soltis et al., 2007	Soltis et al., 2007
Celastrales	2	1	1	1	2	1	pruned	pruned	Soltis et al., 2007	Soltis et al., 2007
Malpighiales	7	6	5	2	7	2	pruned	pruned	Soltis et al., 2007	Soltis et al., 2007
Huaceae	1	1	1	pruned	1	pruned	pruned	pruned	Soltis et al., 2007	Soltis et al., 2007
Oxalidales	2	1	pruned	pruned	2	pruned	pruned	pruned	Soltis et al., 2007	Soltis et al., 2007
Zygophyllaceae	1	pruned	pruned	pruned	1	pruned	pruned	pruned	Soltis et al., 2007	Soltis et al., 2007
Sapindales	2	2	1	1	2	1	pruned	pruned	Soltis et al., 2007	Soltis et al., 2007
Malvales	3	3	3	2	3	2	pruned	pruned	Soltis et al., 2007	Soltis et al., 2007
Brassicales	2	1	1	pruned	2	pruned	pruned	pruned	Soltis et al., 2007	Soltis et al., 2007
Huerteales	1	1	1	pruned	1	pruned	pruned	pruned	Soltis et al., 2007	Soltis et al., 2007
Crossosomatales	3	2	1	pruned	3	pruned	pruned	pruned	Soltis et al., 2007	Soltis et al., 2007
Geraniales	3	3	3	2	3	2	pruned	pruned	Soltis et al., 2007	Soltis et al., 2007
Myrtales	4	4	4	3	4	3	pruned	pruned	Soltis et al., 2007	Soltis et al., 2007
Vitales	2	2	2	1	2	1	pruned	pruned	Soltis et al., 2007	Soltis et al., 2007
Saxifragales	24	19	17	6	24	6	pruned	pruned	Soltis et al., 2007	Soltis et al., 2007; Soltis et al., 1996
Berberidopsidales	2	2	2	1	2	1	pruned	pruned	Soltis et al., 2007	Soltis et al., 2007
Santalales	7	5	5	3	7	3	pruned	pruned	Worberg et al., 2007	Soltis et al., 2007
Asterids:	46	42	36	22	46	22	pruned	pruned	Bremer, B. et al., 2002; Soltis et al., 2007	Bremer, B. et al., 2002; Soltis et al., 2007
Cornales	5	5	5	5	5	5	pruned	pruned	Bremer, B. et al., 2002; Soltis et al., 2007	Bremer, B. et al., 2002; Soltis et al., 2007
Ericales	16	15	14	6	16	6	pruned	pruned	Bremer, B. et al., 2002; Soltis et al., 2007	Bremer, B. et al., 2002; Soltis et al., 2007
Lamiales	8	7	5	3	8	3	pruned	pruned	Bremer, B. et al., 2002; Soltis et al., 2007	Bremer, B. et al., 2002; Soltis et al., 2007
Montiniaceae	1	1	1	1	1	1	pruned	pruned	Bremer, B. et al., 2002	Bremer, B. et al., 2002
Convolvulaceae	1	pruned	pruned	pruned	1	pruned	pruned	pruned	Bremer, B. et al., 2002	Bremer, B. et al., 2002
Vahliaceae	1	1	1	1	1	1	pruned	pruned	Bremer, B. et al., 2002	Bremer, B. et al., 2002
Rubiaceae	1	1	1	pruned	1	pruned	pruned	pruned	Bremer, B. et al., 2002	Bremer, B. et al., 2002
Eucommiaceae	1	1	1	1	1	1	pruned	pruned	Bremer, B. et al., 2002	Bremer, B. et al., 2002
Icacinaceae	2	2	2	1	2	1	pruned	pruned	Bremer, B. et al., 2002	Bremer, B. et al., 2002
Oncothecaceae	1	1	pruned	pruned	1	pruned	pruned	pruned	Bremer, B. et al., 2002	Bremer, B. et al., 2002
Adoxaceae	1	1	pruned	pruned	1	pruned	pruned	pruned	Bremer, B. et al., 2002	Bremer, B. et al., 2002
Paracryphiaceae	1	1	1	pruned	1	pruned	pruned	pruned	Bremer, B. et al., 2002	Bremer, B. et al., 2002
Pittosporaceae	1	1	1	1	1	1	pruned	pruned	Bremer, B. et al., 2002	Bremer, B. et al., 2002
Aralidiaceae	1	1	1	1	1	1	pruned	pruned	Bremer, B. et al., 2002	Bremer, B. et al., 2002
Asterales	2	2	2	1	2	1	pruned	pruned	Bremer, B. et al., 2002	Bremer, B. et al., 2002
Bruniaceae	1	1	1	1	1	1	pruned	pruned	Bremer, B. et al., 2002	Bremer, B. et al., 2002
Tribelaceae	1	1	pruned	pruned	1	pruned	pruned	pruned	Bremer, B. et al., 2002	Bremer, B. et al., 2002
Aquifoliales	1	pruned	pruned	pruned	1	pruned	pruned	pruned	Bremer, B. et al., 2002	Bremer, B. et al., 2002
Dilleniaceae	2	2	1	1	2	1	1	1	Soltis et al., 2007	Soltis et al., 2007
Rhabdodendraceae	1	1	1	1	1	1	1	1	Cuénoud et al., 2002; Soltis et al., 2007	Cuénoud et al., 2002; Soltis et al., 2007
Frankeniaceae	2	1	1	pruned	2	pruned	2	pruned	Soltis et al., 2007	
Tamaricaceae	2	2	1	pruned	2	pruned	2	pruned	Cuénoud et al., 2002; Soltis et al., 2007	
Plumbaginaceae	7	6	4	1	7	1	7	1	Cuénoud et al., 2002; Soltis et al., 2007	
Polygonaceae	8	6	3	1	8	1	8	1	Cuénoud et al., 2002; Soltis et al., 2007	Sanchez & Kron, 2008;
Dioncophyllaceae	1	1	1	pruned	1	pruned	1	pruned	Cuénoud et al., 2002; Meimberg et al., 2000	Cuénoud et al., 2002; Meimberg et al., 2000
Ancistrocladaceae	4	4	2	pruned	4	pruned	4	pruned	Cuénoud et al., 2002; Soltis et al., 2007	Meimberg, et al., 2000
Nepenthaceae	5	5	4	1	5	1	5	1	Cuénoud et al., 2002; Soltis et al., 2007	Meimberg et Heubl, 2006
Droseraceae	7	4	2	1	7	1	7	1	Cuénoud et al., 2002; Soltis et al., 2007	Meimberg, et al., 2000; Cameron et al., 2002; Cuénoud et al., 2002
Simmondsiaceae	1	1	1	1	1	pruned	1	1	Soltis et al., 2007	Soltis et al., 2007
Asteropeiaceae	1	1	1	1	1	1	1	1	Cuénoud et al., 2002; Soltis et al., 2007	Cuénoud et al., 2002; Soltis et al., 2007
Achatocarpaceae	1	1	1	1	1	pruned	1	1	Cuénoud et al., 2002	Cuénoud et al., 2002
Amaranthaceae	10	6	4	2	10	2	10	2	Cuénoud et al., 2002; Müller and Borsch, 2005	Cuénoud et al., 2002, Fior et al., 2006; Müller and Borsch, 2005
Caryophyllaceae	17	12	7	3	17	2	17	3	Fior et al., 2006	Fior et al., 2006; Müller and Borsch, 2005
Molluginaceae	8	6	4	pruned	8	pruned	8	pruned	Cuénoud et al., 2002; Soltis et al., 2007	Cuénoud et al., 2002; Fior et al., 2006
Lophiocarpaceae	1	pruned	pruned	pruned	1	pruned	1	pruned	Cuénoud et al., 2002	Cuénoud et al., 2002
Barbeuiaceae	1	1	1	pruned	1	pruned	1	pruned	Cuénoud et al., 2002	Cuénoud et al., 2002
Phytolaccaceae I	3	3	2	1	3	1	3	1	Cuénoud et al., 2002; Soltis et al., 2007	Cuénoud et al., 2002
Gisekiaceae	1	1	1	1	1	pruned	1	1	Cuénoud et al., 2002	Cuénoud et al., 2002
Aizoaceae	7	4	2	1	7	1	7	1	Soltis et al., 2007	Cuénoud et al., 2002
Phytolaccaceae II	7	5	3	pruned	7	pruned	7	pruned	Cuénoud et al., 2002; Soltis et al., 2007	Cuénoud et al., 2002
Sarcobataceae	1	1	1	1	1	pruned	1	1	Cuénoud et al., 2002	Cuénoud et al., 2002
Nyctaginaceae	9	8	5	1	9	pruned	9	1	Soltis et al., 2007	Douglas et Manos, 2007
Portulacineae:	218	141	62	32	68	6	218	32	Nyffeler & Eggli, 2010a	Nyffeler & Eggli, 2010a
Basellaceae	2	2	1	pruned	2	1	2	pruned	Cuénoud et al., 2002; Nyffeler, 2007	
Halophytaceae	1	pruned	pruned	pruned	1	pruned	1	pruned	Nyffeler, 2007	Nyffeler, 2007
Montiaceae	31	20	12	pruned	31	pruned	31	pruned	Nyffeler & Eggli, 2010b	Nyffeler & Eggli, 2010b, O'Quinn & Hufford, 2005
Didiereaceae	7	5	2	pruned	7	pruned	7	pruned	Nyffeler & Eggli, 2010b	Nyffeler & Eggli, 2010b, Applequist et al., 2006
Talinaceae	7	4	1	pruned	7	pruned	7	pruned	Nyffeler & Eggli, 2010b	Nyffeler & Eggli, 2010b; Applequist et al., 2006
Portulacaceae	3	2	1	1	3	1	3	1	Cuénoud et al., 2002; Applequist et al., 2006; Edwards et al., 2005; Nyffeler, 2007; Wagstaff and Hennion, 2007	Cuénoud et al., 2002; Edwards et al., 2005; Nyffeler, 2007; Wagstaff and Hennion, 2007;
Anacampserotaceae	14	10	1	1	14	1	14	1	Nyffeler, 2007; Nyffeler & Eggli, 2010b	Nyffeler, 2007
Pereskioideae	17	11	5	3	1	1	17	3	Edwards et al, 2005; Nyffeler & Eggli, 2010b	Edwards et al, 2005
Opuntioideae	11	8	2	1	pruned	pruned	11	1	Edwards et al, 2005; Nyffeler & Eggli, 2010b	Edwards et al, 2005
Maihuenioideae	2	2	2	2	pruned	pruned	2	2	Nyffeler & Eggli, 2010b	Edwards et al, 2005
Cactoideae	99	62	26	19	1	1	99	19	Nyffeler, 2002	Nyffeler, 2002
Trichocereinae	24	15	9	5	1	1	24	5	Lendel et al. (unpubl. data)	Lendel et al. (unpubl. data)
total taxa in the data set	460	345	230	115	333	87	300	46		

	dsA	dsB	dsC	dsD	dsE	dsF	dsG	dsH
% of the data set	35.0	40.3	54.3	60.9	48.3	80.5	0.30	2.20
	37.0	35.7	30.4	17.4	51.1	17.2	56.70	43.50
	28.0	24.1	15.2	21.7	0.60	2.30	43.00	54.30

	dsA	dsB	dsC	dsD	dsE	dsF	dsG	dsH
sampled % of the clade biodiversity	0.09	0.07	0.07	0.04	0.09	0.04	0.00	0.00
	1.48	1.07	0.61	0.17	1.48	0.13	1.48	0.17
	6.97	4.49	1.89	1.35	0.11	0.11	6.97	1.35


Supplement S6:
Bayesian tree

Supplement S7:
Strict consensus tree

Supplement S8: Estimated divergence times

data set	method	mixing & converg.	nb of runs	days	CG asterids	SL Caryophyllales	CG Caryophyllales	SL Portulacineae	CG Portulacineae	SL Cactaceae	CG Cactaceae	SL Cactoideae	CG Cactoideae	SL Trichocereinae	CG Trichocereinae
dsA	**UCLN**	well	6x30mil	11	**105.1 (99.5 - 111.1)**	**111.3** (105.9 - 116.4)	**104.8** (98.7- 110.7)	**77.8** (67.5 - 87.6)	**66.3** (55.9 - 77.7)	**47.1** (36.2 - 57.8)	**38.0** (27.4 - 48.7)	**34.0** (24.9 - 43.9)	**29.5** (21.3 - 38.2)	**10.1**	**6.6 (a)** (3.8 - 10.3)
	NPRS	n.a.	n.a.	n.a.	104.0	112.3	100.2	80.1	65.4	58.3	50.6	46.8	42.0	18.9	10.1
	PLFB	n.a.	n.a.	n.a.	x	x	x	x	x	x	x	x	x	x	x
	PLBP	n.a.	n.a.	n.a.	x	x	x	x	x	x	x	x	x	x	x
	PATHd8	n.a.	n.a.	n.a.	89.3	100.4	91.0	70.6	23.1	15.9	11.7	11.6	9.2	2.0	0.8
dsB	**UCLN***	poor	many	8	101.3 (93.0 - 108.4)	105.0 (99.0 - 111.3)	101.8 (95.1 - 108.8)	84.8 (76.2 - 94.6)	81.4 (70.5 - 90.9)	72.0 (60.1 - 83.3)	66.3 (54.1-77.3)	62.0	55.1 (b) (44.0-69.7)	31.2 (c)	27.3 (c) (13.2 - 41.0)
	NPRS	n.a.	n.a.	n.a.	104.0	112.1	100.8	78.6	60.3	53.6	46.7	43.3	38.3	16.4	8.5
	PLFB	n.a.	n.a.	n.a.	x	x	x	x	x	x	x	x	x	x	x
	PLBP	n.a.	n.a.	n.a.	x	x	x	x	x	x	x	x	x	x	x
	PATHd8	n.a.	n.a.	n.a.	89.3	99.8	91.0	70.6	22.4	16.5	12.2	12.2	10.2	2.5	1.0
dsC	**UCLN**	+-well	20x30mil	5	100.0 (94.0 - 106.7)	103.2 (96.4 - 110.0)	99.8 (92.4 - 107.0)	77.2 (64.0 - 91.9)	71.6 (56.0 - 86.1)	58.3 (39.7 - 74.1)	54.2 (36.7 - 71.4)	47.9 (b) (30.6 - 65.5)	40.7 (b) (23.6 - 57.7)	24.0 (c) (11.4 - 38.4)	15.4 (c) (4.9 - 27.7)
	NPRS	n.a.	n.a.	n.a.	104.3	112.1	100.6	74.3	54.5	48.1	40.5	36.8	31.6	12.7	6.6
	PLFB	n.a.	n.a.	n.a.	x	x	x	x	x	x	x	x	x	x	x
	PLBP	n.a.	n.a.	n.a.	x	x	x	x	x	x	x	x	x	x	x
	PATHd8	n.a.	n.a.	n.a.	89.3	102.1	91.0	70.6	24.0	16.1	12.8	12.7	10.6	2.6	1.3
dsD	**UCLN**	well	16x30mil	6	100.5 (93.6 - 108.5)	103.6	97.4 (90.1 - 106.9)	79.8	71.1 (51.3 - 89.9)	71.1 (51.3 - 89.9)	63.1 (41.5 - 83.7)	56.4 (b) (32.8 - 77.5)	45.7 (b) (22.9 - 68.0)	23.5 (6.8 - 40.5)	12.9 (1.6 - 27.2)
	NPRS	n.a.	n.a.	n.a.	103.8	111.3	98.5	79.7	56.1	56.1	48.5	44.0	38.0	14.4	8.2
	PLFB	n.a.	n.a.	n.a.	106.4	115.0	100.2	80.6	53.6	53.6	46.3	42.4	36.9	16.3	10.1
	PLBP	n.a.	n.a.	n.a.	108.4	117.0	101.7	81.2	43.3	43.3	26.4	24.1	20.8	5.3	2.7
	PATHd8	n.a.	n.a.	n.a.	89.3	99.3	91.0	70.6	17.3	16.8	13.2	13.1	11.3	2.8	1.6
dsE	**UCLN**	+-well	20x30mil	6	99.1 (94.0 - 104.7)	101.9 (96.8 - 107.9)	99.3 (93.5 - 105.5)	80.3 (67.3 - 92.3)	74.5 (60.3 - 87.9)	46.1 (29.3 - 63.6)	19.1 (0.8 - 43.4)	pruned	pruned	pruned	pruned
	NPRS	n.a.	n.a.	n.a.	103.7	111.8	101.0	80.0	65.3	58.3	48.3	pruned	pruned	pruned	pruned
	PLFB	n.a.	n.a.	n.a.	x	x	x	x	x	x	x	x	x	x	x
	PLBP	n.a.	n.a.	n.a.	x	x	x	x	x	x	x	x	x	x	x
	PATHd8	n.a.	n.a.	n.a.	89.3	102.6	91.0	70.6	31.1	24.1	7.1	pruned	pruned	pruned	pruned
dsF	**UCLN**	well	3x30mil	3	104.4 (98.1 - 110.5)	109.7 (104.0 - 115.7)	100.6 (92.8 - 108.5)	77.7 (70.4 - 85.9)	37.4 (15.6 - 61.9)	32.5 (12.8 - 56.0)	17.5 (2.7 - 37.0)	pruned	pruned	pruned	pruned
	NPRS	n.a.	n.a.	n.a.	103.8	111.2	97.5	80.2	55.9	53.3	43.9	pruned	pruned	pruned	pruned
	PLFB	n.a.	n.a.	n.a.	103.5	111.1	97.4	80.1	51.5	48.8	39.7	pruned	pruned	pruned	pruned
	PLBP	n.a.	n.a.	n.a.	108.3	116.8	101.4	81.0	44.6	41.8	27.3	pruned	pruned	pruned	pruned
	PATHd8	n.a.	n.a.	n.a.	89.3	101.3	93.1	70.6	23.0	16.8	8.2	pruned	pruned	pruned	pruned
dsG	**UCLN**	well	4x30mil	6	pruned	n.a.	100 (fixed)	65.4 (53.1 - 76.5)	53.5 (42 - 65.7)	36.6 (27.7 - 46.2)	30.5 (22.2 - 39.5)	27.7 (20.1 - 36.2)	24.5 (16.9 - 31.8)	7.8 (a)	6.4 (a) (3.4 - 9.8)
	NPRS	n.a.	n.a.	n.a.	pruned	n.a.	100 (fixed)	80.6	66.5	60.1	52.5	48.4	43.8	19.4	10.2
	PLFB	n.a.	n.a.	n.a.	pruned	n.a.	100 (fixed)	87.1	74.4	68.2	60.9	57.4	52.2	27.1	10.8
	PLBP	n.a.	n.a.	n.a.	pruned	x	100 (fixed)	x	x	x	x	x	x	x	x
	PATHd8	n.a.	n.a.	n.a.	pruned	n.a.	100 (fixed)	70.6	28.7	20.8	15.1	14.9	12.1	2.6	1.1
dsH	**UCLN**	well	3x30mil	2	pruned	n.a.	100 (fixed)	79.8	69.4 (48.4 - 91)	69.4 (48.5-91)	62 (38.8-83.2)	54.8 (b) (31.6-77.6)	45.4 (b) (23.8-67.5)	19.4	13.3 (2.5 - 26.8)
	NPRS	n.a.	n.a.	n.a.	pruned	n.a.	100 (fixed)	79.3	55.8	55.8	47.8	43.8	38.1	15.5	8.7
	PLFB	n.a.	n.a.	n.a.	pruned	n.a.	100 (fixed)	79.3	54.4	54.4	45.6	41.6	35.8	14.0	7.8
	PLBP	n.a.	n.a.	n.a.	pruned	n.a.	100 (fixed)	80.4	53.3	53.3	45.9	42.9	38.2	20.1	13.3
	PATHd8	n.a.	n.a.	n.a.	pruned	n.a.	100 (fixed)	70.6	21.3	20.7	16.5	16.2	13.8	3.2	1.9
dsI	**UCLN**	well	10x30mil	14	111 (101.3 - 121.0)	119.8 (108.3 - 130.8)	111.0 (101.0 - 121.4)	80.5 (69.3 - 91.7)	67.6 (55.9 - 79.4)	47.6 (37.2 - 58.8)	38.3 (28.0 - 49.8)	34.5 (25.4- 45.1)	29.8 (21.1 - 38.8)	7.3 (a)	6.7 (a) (3.7 - 10.0)
	NPRS	n.a.	n.a.	n.a.	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)
	PLFB	n.a.	n.a.	n.a.	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)
	PLBP	n.a.	n.a.	n.a.	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)
	PATHd8	n.a.	n.a.	n.a.	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)
dsJ	**UCLN**	+-well	17x30mil	3	100.3 (92.7 - 109.1)	103.5	97.0 (90.1 - 107.0)	81.8	68.6 (48.6 - 87.2)	68.6 (48.6 - 87.2)	61.6 (40.9 - 81.6)	54.6 (b) (34.4 - 74.6)	44.9 (b) (23.7 - 67.5)	22.9 (6.3 - 38.8)	12.5 (1.2 - 26.0)
	NPRS	n.a.	n.a.	n.a.	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)
	PLFB	n.a.	n.a.	n.a.	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)
	PLBP	n.a.	n.a.	n.a.	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)
	PATHd8	n.a.	n.a.	n.a.	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)	not applicable (1)
dsK	**UCLN**	well	8x30mil	10	105.1 (99.5 - 111.0)	111.0 (105.2 - 116.7)	104.4 (98.2 - 111.0)	75.3 (64.9 - 86.3)	63.8 (52.4 - 74.7)	45.4 (34.4 - 56.0)	37.0 (27.1 - 47.8)	33.2 (24.3 - 43.2)	28.6 (24.3 - 43.1)	8.9	6.6 (a) (3.8 - 9.7)
	NPRS	n.a.	n.a.	n.a.	103.9	112.3	100.2	79.9	65.2	58.1	50.4	46.6	41.9	18.9	10.1
	PLFB	n.a.	n.a.	n.a.	x	x	x	x	x	x	x	x	x	x	x
	PLBP	n.a.	n.a.	n.a.	x	x	x	x	x	x	x	x	x	x	x
	PATHd8	n.a.	n.a.	n.a.	89.3	100.0	91.0	39.3	23.1	15.9	11.7	11.6	9.2	2.0	0.8
dsL	**UCLN**	well	13x30mil	4	99.3 (92.4 - 107.1)	101.5	95.4 (90.1 - 104.3)	72.0	63.6 (40.2 - 87.3)	63.6 (40.2 - 87.3)	56.9 (32.8 - 78.6)	50.6 (b) (28.0 - 73.3)	41.7 (b) (20.2 - 63.2)	18.1	12.0 (1.5 - 25.0)
	NPRS	n.a.	n.a.	n.a.	103.7	111.2	97.9	68.2	45.1	45.1	38.6	34.9	30.2	11.5	6.5
	PLFB	n.a.	n.a.	n.a.	104.1	111.9	98.0	68.0	44.9	44.9	38.4	34.8	30.1	11.6	6.6
	PLBP	n.a.	n.a.	n.a.	108.1	116.6	98.8	66.6	34.1	34.1	19.3	17.8	15.4	4.0	2.1
	PATHd8	n.a.	n.a.	n.a.	89.3	98.3	91.0	39.2	17.3	16.8	13.2	13.1	11.3	2.8	1.6
	Average age (SD)				**104.2 (+-2.9)**	**110.8 (+- 5.0)**	**100.2 (+-3.3)**	**77.8 (+- 5.0)**	**58.4 (+- 10.9)**	**52.5 (+- 10.1)**	**43.2 (+-12.7)**	**41.1 (+- 10.1)**	**35.6 (+-8.7)**	**15.3 (+-6.2)**	**9.1 (+-3.4)**
						<=== fossil constrained		not constrained ===>							

data set	method	mixing & converg.	nb of runs	days	CG asterids	SL Caryophyllales	CG Caryophyllales	SL Portulacineae	CG Portulacineae	SL Cactacaeae	CG Cactaceae	SL Cactoideae	CG Cactoideae	SL Trichocereinae	CG Trichocereinae
dsM	UCLN	+-well	10x30mil	13	76.9 (60.8 - 92.3)	93.7 (80.9 - 104.9)	86.5 (74.4 - 98.2)	58.3 (47.3 - 69.6)	47.0 (36.5 - 57.6)	32.8 (24.5 - 42.0)	25.7 (18.4 - 33.1)	23.1 (16.5 - 29.5)	19.8 (14.3 - 25.8)	6.8	4.7
	NPRS	n.a.	n.a.	n.a.	95.6	105.3	92.6	73.9	61.7	55.9	48.9	45.4	41.2	19.2	10.2
	PLFB	n.a.	n.a.	n.a.	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)
	PLBP	n.a.	n.a.	n.a.	x	x	x	x	x	x	x	x	x	x	x
	PATHd8	n.a.	n.a.	n.a.	60.8	99.6	76.3	39.3	23.1	15.9	11.7	11.6	9.2	2.0	0.8
dsN	UCLN	+-well	9x30mil	3	50.3 (26.6 - 76.2)	60.3 (33.2 - 88.1)	52.7 (26.9 - 80.7)	37.1	33.3 (15.9 - 52.3)	33.3 (15.9 - 52.3)	29.5 (13.0 - 45.4)	26.4 (b) (11.1 - 42.5)	21.2 (b) (9.0 - 35.4)	8.6	5.7 (0.6 - 12.3)
	NPRS	n.a.	n.a.	n.a.	79.5	99.7	71.7	48.8	32.1	32.1	27.4	24.8	21.4	8.1	4.6
	PLFB	n.a.	n.a.	n.a.	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)
	PLBP	n.a.	n.a.	n.a.	90.0	104.2	84.9	55.4	27.0	27.0	14.5	13.4	11.7	3.1	1.6
	PATHd8	n.a.	n.a.	n.a.	67.9	98.0	77.6	39.2	17.3	16.8	13.2	13.1	11.3	2.8	1.6
dsO	UCLN*	poor	many	7	30.4 (17.1 - 64.3)	33.7 (19.1 - 64.9)	32.5 (18.6 - 64.2)	23.1 (13.6 - 47.4)	22.1 (12.9-47.1)	14.0 (6.4-32.4)	5.3 (0.1 - 15.2)	pruned	pruned	pruned	pruned
	NPRS	n.a.	n.a.	n.a.	93.1	103.6	91.3	72.1	58.7	52.2	43.0	pruned	pruned	pruned	pruned
	PLFB	n.a.	n.a.	n.a.	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)
	PLBP	n.a.	n.a.	n.a.	x	x	x	x	x	x	x	pruned	pruned	pruned	pruned
	PATHd8	n.a.	n.a.	n.a.	89.3	102.6	91.0	70.1	31.1	24.1	7.1	pruned	pruned	pruned	pruned
dsP	**UCLN**	well	2x30mil	3	75.8 (60.4 - 90.3)	87.3 (72.9 - 100.8)	74.6 (89.5 - 69.5)	46.1 (31.9 - 60.8)	25.3 (14.6 - 37.8)	22.9 (12.1 - 33.7)	12.8 (3.4 - 23.4)	pruned	pruned	pruned	pruned
	NPRS	n.a.	n.a.	n.a.	77.1	89.1	67.6	44.3	24.1	22.8	18.4	pruned	pruned	pruned	pruned
	PLFB	n.a.	n.a.	n.a.	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)
	PLBP	n.a.	n.a.	n.a.	89.4	104.0	86.5	56.8	29.1	27.0	15.6	pruned	pruned	pruned	pruned
	PATHd8	n.a.	n.a.	n.a.	70.7	99.4	90.4	44.1	23.0	16.8	8.2	pruned	pruned	pruned	pruned
dsQ	UCLN	well	3x30mil	6	pruned	n.a.	100 (fixed)	57.0 (44.2 - 70.8)	46.1 (34.6 - 59.0)	32.6 (22.5 - 41.4)	26.3 (18.1 - 34.7)	23.5 (16.6 - 31.1)	20.5 (13.9 - 27.3)	5.3	4.9 (a) (2.6 - 7.4)
	NPRS	n.a.	n.a.	n.a.	pruned	n.a.	100 (fixed)	79.3	65.5	59.3	51.8	47.7	43.1	19.2	10.0
	PLFB	n.a.	n.a.	n.a.	pruned	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)
	PLBP	n.a.	n.a.	n.a.	pruned	x	100 (fixed)	x	x	x	x	x	x	x	x
	PATHd8	n.a.	n.a.	n.a.	pruned	n.a.	100 (fixed)	49.6	28.7	20.5	15.1	14.9	12.1	2.6	1.1
dsR	**UCLN**	well	4x30mil	2	pruned	n.a.	100 (fixed)	59.5	53.5 (27.9 - 78.6)	53.5 (27.9 - 78.6)	47.1 (24.3 - 72.4)	41.4 (20.4 - 64.6)	34.2 (b) (15.8 - 55.4)	17.9 (6.0 - 31.8)	9.7 (1.4 - 20.6)
	NPRS	n.a.	n.a.	n.a.	pruned	n.a.	100 (fixed)	65.0	42.7	42.7	36.1	33.1	28.8	11.7	6.6
	PLFB	n.a.	n.a.	n.a.	pruned	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)	not applicable (2)
	PLBP	n.a.	n.a.	n.a.	pruned	n.a.	100 (fixed)	69.5	45.9	45.9	39.4	36.8	32.7	16.8	11.1
	PATHd8	n.a.	n.a.	n.a.	pruned	n.a.	100 (fixed)	47.9	21.3	20.7	16.5	16.2	13.8	3.2	1.9
		<=== fossil constrained			not constrained ===>										

Legend:

SL = stem lineage
CG = crown group

UCLN = an uncorrelated lognormal relaxed-clock model, with confidence interval values in brackets
UCLN* = poor converging and/or mixing of the MCC chains
NPRS = nonparametric rate smoothing method
PLFB = penalized likelihood with fossils-based rate smoothing
PLBP = penalized likelihood with branch-pruning rate smoothing

n.a. = not available
days = time counted in days, needed to finish 1 analysis of 30 mil. generations
x = cross validation failed
pruned = excluded from the analysis

not applicable(1) = requires at least one fixed age
not applicable(2) = requires at least one fixed and two constrained ages

(a) = excluding *Cleistocactus icosagonus* and *Cleistocactus sepium*
(b) = excluding *Blossfeldia liliputana*
(c) = excluding *Cleistocactus icosagonus*

Supplement S9a: Comparison of the topologies

clade	fossil	fossil age (BEAST)	Selected tree	Parsimony tree	Bayesian tree	UCLN tree
CGEudicots	1	125	ok	ok	ok	ok
SL Menispermaceae	2	65.5	ok	ok	ok	ok
SL Berberidaceae	3	33.9	ok	ok	ok	ok
Ranunculales	-	-	monophyletic and at the base of the tree	not monophyletic, diverging first at the base of the core eudicots	monoph.,unresolved with Sabiales, Proteales, and the rest of basal eudicots	momoph.,sister to (Sabiales +Nelumbon.)
SL Nelumbonaceae	4	105.8 (103)	ok	ok	ok	not well defined, falls on the same node as fossil 6
SL Proteaceae	5	85 (83.5)	ok	ok	ok	ok
SL Sabiaceae	6	98 (93.5)	ok	owerwrites the fossil No 4	ok	not well defined, falls on the same node as fossil 4
Sabiales	-	-	monoph., unresolved with Proteales and (Buxales+the rest)	monoph., sister to (Proteales+(Trochod+the rest))	monoph., unresolved with Proteales, Ranunculales and and the rest of basal eudicots	monoph., sister to Nelumbonaceae; the two together are sister to Ranunculales
Proteales	-	-	monoph., unresolved with Sabiales and (Buxales+the rest)	monoph., sister to (Trochodend+the rest)	monoph., unresolved with Sabiales, Ranunculales and and the rest of basal eudicots	polyphyletic, in a clade with Sabiales and Ranunculales, all together sister to the rest of the basal eudicots
SL Buxales	7	99.6	ok	at the SL of Trochodendrales (or at the SL Buxales+core eudicots)	at the SL of Trochodendrales (or at the SL Buxales+core eudicots)	ok
SL Buxaceae	8	94	ok	ok	ok	ok
Buxales	-	-	unresolved with Trochodendrales and core eudicots	sister to core eudicots	sister to core eudicots	sister to (Trochodendrales + core eudicots)
Trochodendrales	-	-	unresolved with Buxales and core eudicots	sister to (Buxales + core eudicots)	sister to (Buxales + core eudicots)	sister to core eudicots
core eudicots	-	-	monphyletic, unresolved with Buxales and Trochodendrales	monphyletic, sister to Buxales	monphyletic, sister to Buxales	monphyletic, sister to Trochodendrales
Gunnerales	-	-	monphy., basal to core eudicots	monphy., basal to core eudicots	monphy., basal to core eudicots	monphy., basal to core eudicots
SL Gunneraceae	9	91	ok	ok	ok	ok
eurosids I	-	-	((not monoph., unresolved with Zygophyllaceae)+ (eurosids II+ Crossomatales)+Geraniales+Myrtales)	not monophyletic, unresolved with (eurosids II+Myrtales), Geraniales, Crossomatales, Zygophyllales	(form clade with Zygophyllales) unresolved with (eurosidsII+Myrtales), crossomatales and Geraniales	(form clade with Zygophyllales) sister to a clade of (eurosidsII, Myrtales, Crossomatales, Geraniales)
CG Fagales	10	93.5	ok	ok	ok	ok
SL Juglandaceae	11	83.5 (77)	ok	overwrites fossil 10	overwrites fossil 10	includes one node from above!
SL Rhamnaceae	12	48.6	ok	ok	ok	ok
SL Ulmaceae	13	55.8	ok	ok	ok	ok
CG Malpighiales	14	89.3	ok	ok	ok	ok
eurosids II	-	-	sister to Crossomatales	sister to Myrtales	sister to Myrtales	sister to Myrtales
SL Sapindales	15	55.8	ok	ok	ok	ok
CG Malvales	16	33.9	ok	ok	ok	ok
Vitales	-	-	sister to eurosids	sister to Dilleniales; the two unresolved with Saxifragales and eurosids	sister to eurosids	sister to Dilleniales, the two sister to Saxifragales
SL Saxifragales	17	91 (89.3)	ok	ok	ok	ok
Saxifragales	-	-	sister to Vitales (eurosids)	unresolved with (Vitales+Dilleniales), eurosids, Berbidales(rest of the tree)	sister to Vitales(eurosids)	sister to (Vitales+Dilleniales), alltogether sister to eurosids
CG Hamamelidaceae	18	83.5	ok	not monophyletic	ok	ok
SL Altingiaceae	19	89.3	ok	overwrites Fossil 18	overwrites Fossil 18 and 17 ?	includes one node from above! And falls on the same node as the fossil 19
basal core eudicots	-	-	(Saxcifragales + eurosids) (Berberidopsidales (rest of the core eudicots))	unresolved:eurosids, Saxifragales, (Vitales+Dilleniales), Berberidopsidales (rest of the core eudicots)	(Saxcifragales + eurosids) (Berberidopsidales (rest of the core eudicots))	((Vitales+Dilleniales) + Saxifragales) + eurosids sister to Santalales (Berberidopsidales(asterids + Caryophyllales)
Berberidopsidales	-	-	sister clade to a grade of Santalales, asterids, Dilleniales and Caryophyllales	sister clade to a grade of Santalales, asterids and Caryophyllales	sister clade to Santalales (asterids + Caryophyllales)	sister to (asterids + Caryophyllales)

clade	fossil	fossil age (BEAST)	Selected tree	Parsimony tree	Bayesian tree	UCLN tree
Santalales	-	-	unresloved with asterids, Dilleniales and Caryophyllales	unresolved with asterids and Caryophyllales	sister to (asterids+Caryophyllales)	sister to Berberidopsidales (asterids+Caryophyllales)
asterids	-	-	unresloved with Santalales Dilleniales and Caryophyllales	unresolved with Santalales and Caryophyllales	sister to Caryophyllales	sister to Caryophyllales
Dilleniales	-	-	unresolved with Santalales, asterids and Caryophyllales	sister to Vitales; together unresolved with Saxifragales, a clade of eurosids, Crossomatales and Myrtales; and Berberidopsidales(Santalales(Caryophyllales)	unresolved with (saxifragales+eudicots) and Berberidopsidales(Santalales(Caryophyllales)	sister to Vitales; together sister to Saxifragales, all together sistr to eudicots
SL Iteaceae	20	89.3	ok	ok	ok	ok
SL Hydrangeaceae	21	89.3	ok	ok	ok	ok
Cornales	-	-	basal to asterids; sister to Ericales+(euasteridsl+euasteridsll)	unresolved with Ericales and (euasteridsl+euasteridsll)	sister to (euasteridsl+euasteridsll)	sister to (euasteridsl+euasteridsll)
SL Cornaceae	22	87 (85.5)	ok	ok	ok	ok
CG Cornaceae	23	55.8	ok	ok	ok	ok
Ericales	-	-	sister to (euasteridsl +euasteridsll)	unresolved with Cornales and (euasteridsl+euasteridsll)	basal to asterids; sister to Cornales+(euasteridsl+euasteridsll)	basal to asterids; sister to Cornales+(euasteridsl+euasteridsll)
CG Ericales	24	89.3	ok	ok	ok	ok
SLPentaphylacaceae	25	65.5	ok	ok	ok	ok
euasterids I and II	-	-	sister clades, together sister to Ericales	sister clades; unresolved with Cornales and Ericales	sister clades, together sister to Cornales	sister clades, together sister to Cornales
SL Vahliaceae	26	83.5(77)	ok	ok	(ok?: SLVahliales+Gentianales)	not well defined: SLVahliales + Gentianales + Solanales)
CG Apiales	27	37.2	ok	ok	ok	ok
Caryophyllales	-	-	unresolved with Dilleniales, asterids and Santalales	unresolved with asterids and Santalales	sister to asterids	sister to asterids
Rhabdodendron	-	-	basal to Caryophyllales	sister to Simmondsia, not basal to Caryophyllales	not basal to Caryophyllales, unresolved with Simmondsia and Asteropeia(rest of Caryophyllales)	not basal to Caryophyllales (but to Caryophyllales excluding carnivorous plants, Plumbaginaceae and Ploygonaceae)
Frankeniaceae + Tamaricaceae	-	-	unresolved with Plumbaginaceae and carnivorous clade	sister to (Plumbaginaceae+Polygonaceae); all together at the base of Caryophyllaes	sister to (Plumbaginaceae+Polygonaceae); all together sister to the carnivorous calde	sister to Plumbaginaceae+Polygonaceae; all together sister to the carnivorous calde
Plumbaginaceae + Polygonaceae	-	-	unresolved with (Frankeniaceae + Tamaricaceae) and carnivorous clade	sister to (Frankeniaceae + Tamaricaceae); all together at the base of Caryophyllaes	sister to (Frankeniaceae + Tamaricaceae); all together sister to the carnivorous calde	sister to (Frankeniaceae + Tamaricaceae); all together sister to the carnivorous calde
SL Polygonaceae	28	5.33	ok	ok	ok	ok
Carnivorous clade + Plumbaginaceae	-	-	unresolved with (Frankeniaceae + Tamaricaceae) and(Plumbaginaceae+Polygonaceae)	sister to the(Rhabodedndron +rest of Caryophyllales)	sister to ((Frankeniaceae + Tamaricaceae) + (Plumbaginaceae+Polygonaceae))	sister to ((Frankeniaceae + Tamaricaceae) + (Plumbaginaceae+Polygonaceae))
Droseraceae	-	-	sister to Nepenthes	basal to carnivorous clade	basal to carnivorous clade	basal to carnivorous clade
SL Droseraceae	29	91 (89.3)	ok	ok	ok	ok
Nepenthaceae	-	-	sister to Droseraceae	sister to Ancistrocladaceae	sister to Ancistrocladaceae	sister to Ancistrocladaceae
Ancistrocladaceae	-	-	sister to (Nepenthaceae + Droseraceae)	sister to Nepenthaceae	sister to Nepenthaceae	sister toTriphyophyllum peltatum
Carnivorous clade			(Droseraceae+Nepenthaceae) + (Drosophyllum lusitanicum(Triphyophyllum peltatum(Ancistrocaldus)	grade: Droseraceae(Nepenthaceae(Drosophyllum lusitanicum(Triphyophyllum peltatum(Ancistrocaldus)	grade: Droseraceae(Nepenthaceae(Drosophyllum lusitanicum(Triphyophyllum peltatum(Ancistrocaldus)	grade: Droseraceae(Nepenthaceae(Drosophyllum lusitanicum(Triphyophyllum peltatum(Ancistrocaldus)
Simmondsia	-	-	sister to Asteropeiacea(rest of the Caryophyllales)	sister to Rhabdodendron	unresolved with Rhabdodendron and the rest of the Caryophyllales	sister to Asteropeiacea(rest of tha Caryophyllales)
Asteropeiaceae	-	-	sister to the rest of the Caryophyllales	sister to the rest of the Caryophyllales	sister to the rest of the Caryophyllales	sister to the rest of the Caryophyllales

clade	fossil	fossil age (BEAST)	Selected tree	Parsimony tree	Bayesian tree	UCLN tree
Amaranthaceae	-	-	sister to Caryophyllaceae	in a grade: Caryophyllaceae(Amaranthaceae(rest of the Caryophyllales)	sister to Caryophyllaceae	sister to Caryophyllaceae
CG Chenopodioideae / Amaranthaceae	30	55.8	ok	ok	ok	ok
Caryophyllaceae	-	-	sister to Amaranthaceae	in a grade: Asteropeicaeae(Caryophyllaceae(Amaranthaceae(rest of the Caryophyllales))	sister to Amaranthaceae	sister to Amaranthaceae
SL higher Caryophyllaceae	31	33.9	actually CG ??	actually CG ??	actually CG ??	actually CG ??
Limeum	-	-	unresolved with Molluginaceae, clade of Phytolaccaceae, Aizoaceae, Nyctaginaceae; and Portulacineae	in a grade: (Amaranthaceae((Aizoaceae; Phytolaccacea; Nyctaginaceae)(Limeum(Molluginaceae(Portulacinaeae))))	basal to a clade of Aizoacaeae, Nyctaginaceae and Phytolaccaceae	basal to a clade of Aizoacaeae, Nyctaginaceae and Phytolaccaceae
Molluginaceae	-	-	unresolved with Limeum clade of Phytolaccaceae, Aizoaceae, Nyctaginaceae; and Portulacineae	sister to Portulacineae	sister to Portulacineae	sister to Portulacineae
Phytolaccaceae	-	-	in a calde wit Nyctaginaceae and Aizoaceae, all together sister to Portulacineae	in a calde wit Nyctaginaceae and Aizoaceae, all together sister to Molluginaceae(Portulacinaeae)	in a calde wit Nyctaginaceae and Aizoaceae, all together sister to Molluginaceae(Portulacinaeae)	in a calde wit Nyctaginaceae and Aizoaceae, all together sister to Molluginaceae(Portulacinaeae)
SL Phytolaccaceae s.str.	32	70.6	ok	ok	ok	ok
Portulacineae	-	-	sister to a clade of Phytolaccaceae, Aizoaceae and Nyctaginaceae	sister to Molluginaceae	sister to Molluginaceae	sister to Molluginaceae
Montiaceae	-	-	at the base of the Portulacineae, unresolved with Basellaceae and Halophyttaceae	at the base of the Portulacineae	at the base of the Portulacineae	at the base of the Portulacineae
Basellaceae	-	-	unresolved with Halophytaceae, Montiaceae and a grade of Didieraceae(...)	clustered with Ceraria and Portulacaria (Didieriaceae), both unresolved within Portulacinaeae	clustered with Ceraria and Portulacaria (Didieriaceae), both unresolved within Portulacinaeae	clastered with Didieriaceae sister to Tallinaceae (rest of the tree)
Halophytaceae	-	-	unresolved with Basellaceaae, Montiaceae and a grade of Didieraceae(...)	unresolved within Portulacinaeae	unresolved within Portulacinaeae	grade: Montiaceae(Halophytaceae(Tallinaceae(...)
Anacampserotaceae	-	-	sister to Cactaceae	sister to Cactaceae	sister to Cactaceae	sister to Cactaceae
Pereskia	-	-	not monophyletic, a grade of few taxa at the base of Cactaceae	not monophyletic, a grade of few taxa at the base of Cactaceae	not monophyletic, a grade of few taxa at the base of Cactaceae	not monophyletic, at the base of Cactaceae
Cactoideae	-	-	monophyletic with Blossfeldia at the base	monophyletic with Blossfeldia at the base	monophyletic with Blossfeldia at the base	monophyletic with Blossfeldia at the base
Trichocereinae	-	-	monophyletic and unresolved within Cereeinae	not monophyletic (Cle_ico; Cle_sep); unresolved within Cereeinae	paraphyletic (incledes Facheiroa); unresolved within Cereeinae	polyphyletic (Cle_ico; Cle_sep not included); paraphyletic (including Facheiroa) not much support within Cereeinae

Supplement S9b: *Comparison of the Topologies*

We compared the four topologies (i.e., Selected tree (Fig. 2), MP strict consensus tree (Supp. S7), Bayesian consensus phylogram (Supp. S6), and Bayesian maximum clade credibility (MCC) chronogram derived from the BEAST analyses (Fig. 3) as based on the exhaustive data set A for the presence of more than 60 specific clades and sister-group relationships. Overall, inferred relationships were very similar, and depicted relationships in the topologies derived from the different phylogenetic analyses are mainly congruent to the composite topology.

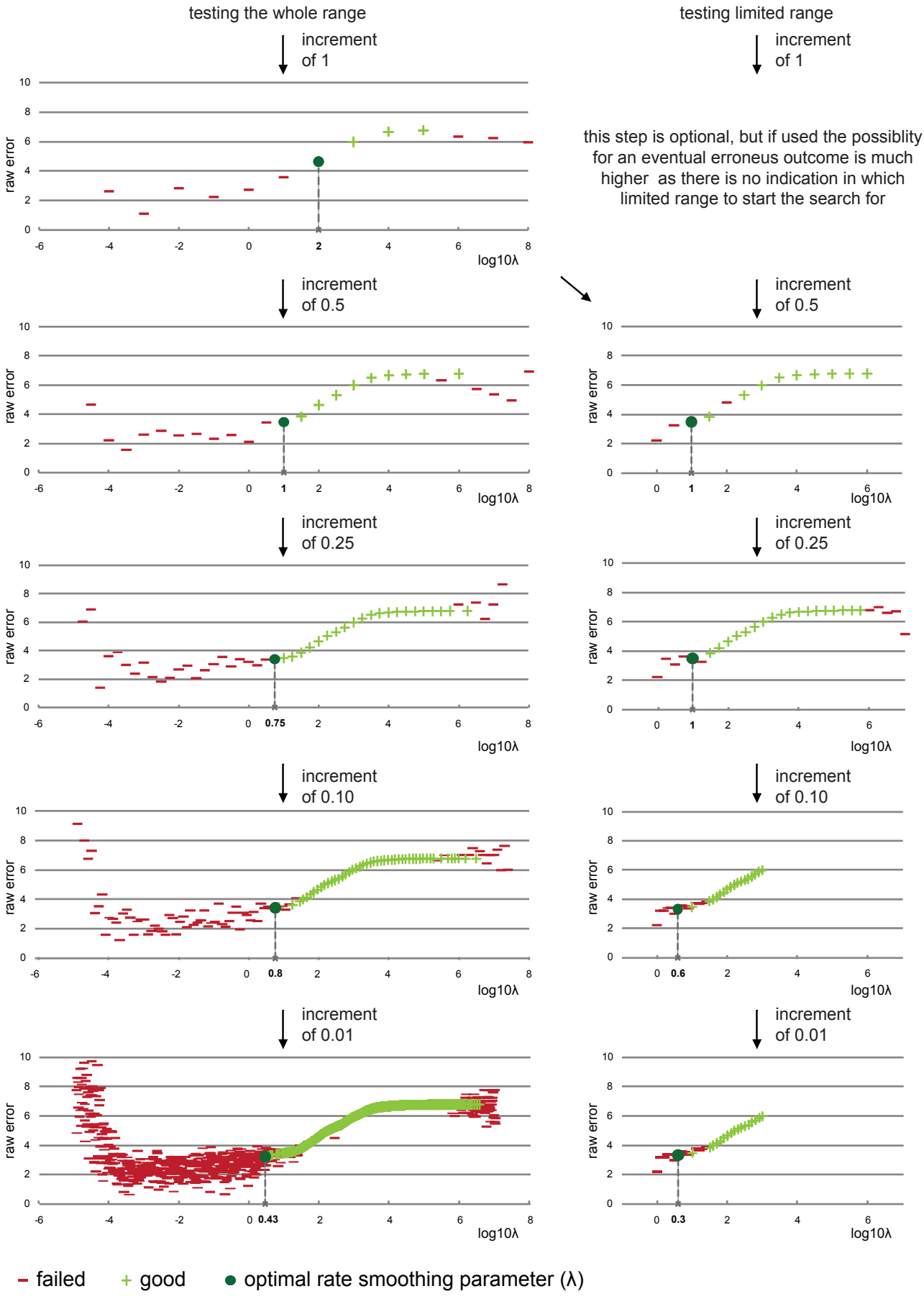
Furthermore, we inspected the effect of the inferred relationships on the redundancy and obsolescence of all fossil constraints. In particular, we verified the topology and the placement of the fossils in the MCC chronogram, as neither the topology of the MP strict consensus tree nor the topology of the Bayesian consensus phylogram were further considered in the divergence time estimations.

The order Ranunculales, which is often inferred as the earliest diverging extant lineage of eudicots, was found to be sister to the remaining eudicot representatives in all analyses expect for UCLN, where it is sister to Proteales+Sabiales. The only case of obsolescence of a fossil constraint was found in this case, where a fossil of stem Nelumbonaceae applies to the same node as the stem lineage of Sabiaceae. The MP strict consensus tree resolves further relationships within the basal eudicots, showing a grade formed by Sabiales, Proteales, Trochodendrales and Buxales; Bayesian consensus phylogram depicts Trochodendrales to be the sister-group to a clade consisting Buxales and the core eudicot taxa, while the MCC chronogram found Buxales to diverge prior to Trochodendrales. Gunnerales is found in all topologies basal within the core eudicots. The order Saxifragales clustered with Vitales and Dilleniales, all together forming the sister-group to the Rosids in the MCC chronogram. The Bayesian

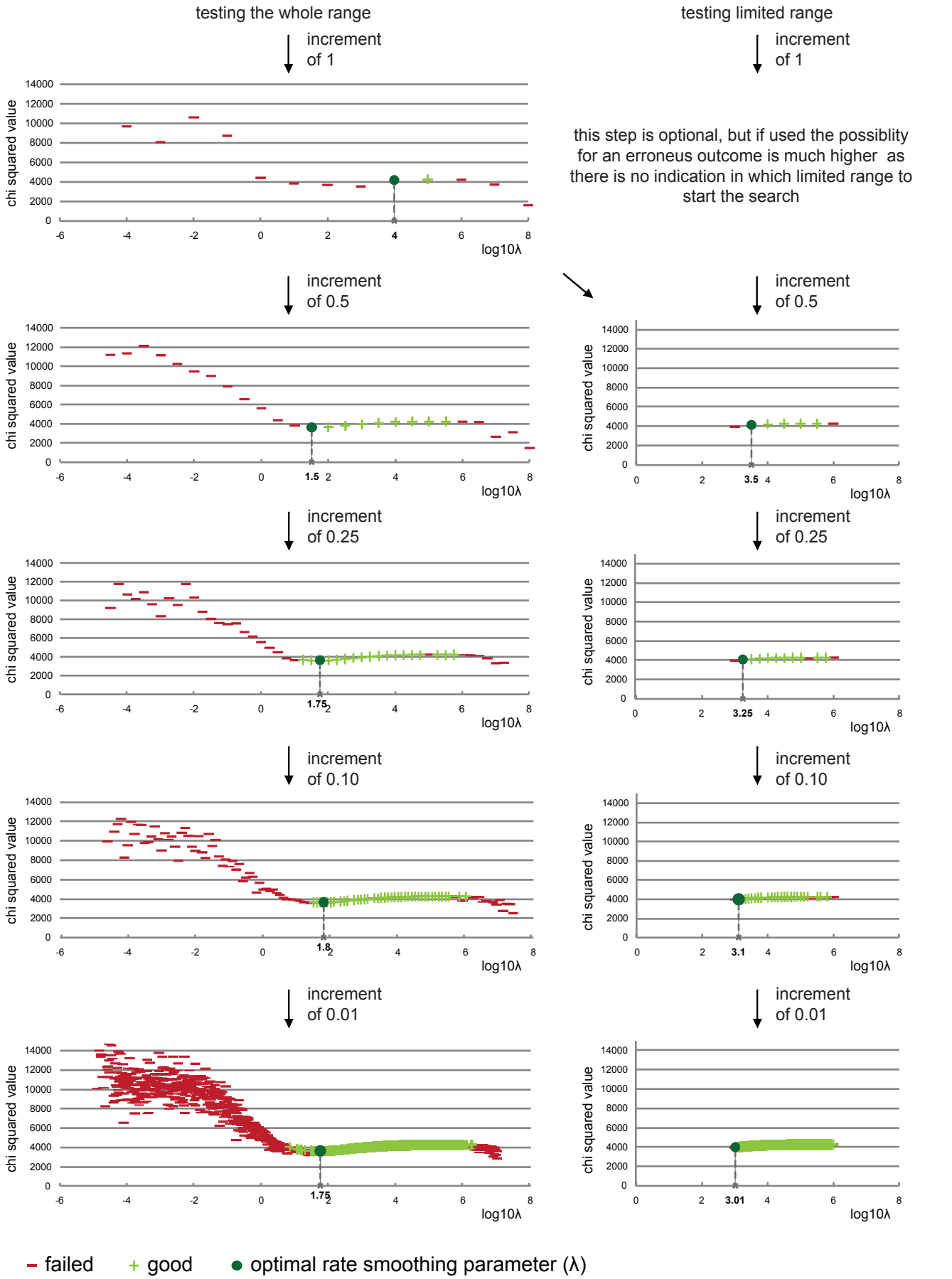
consensus phylogram depicts, similarly to the topology of the designed tree, Saxifragales to be the sister group to the clade of Vitales and Rosids together forming a polytomy with Dilleniales and Berberidopsidales (i.e., Berberidopsidaceae and Aextoxicaceae) that is sister to the remaining eudicots. The clade of Berberidopsidales, Santalales, asterids and Caryophyllales is inferred in all topologies, however in a slightly different relationship: Berberidopsidales is found to diverge first followed by an unresolved divergence of Santalales, asterids and Caryophyllales in the MP strict consensus tree, or excluding the order Santalales forming a sister-group to asterids and Caryophyllales in the Bayesian consensus phylogram. In the MCC chronogram, the basal position was occupied by Santalales followed by a split between Berberidopsidales and asterids sister to Caryophyllales. Sister-group relationship between asterids and Caryophyllales was recovered in both Bayesian consensus phylogram and the MCC chronogram. The clade of Frankeniaceae sister to Tamaricaceae, and Plumbaginaceae sister to Polygonaceae in a basal position within the Caryophyllales is inferred by the MP strict consensus tree, while in the Bayesian consensus phylogram and the MCC chronogram this clade is most closely related to a clade formed by Ancistrocladaceae sister to the carnivorous Caryophyllales taxa, all forming an early diverging clade in Caryophyllales. *Rhabdodendron macrophyllum* is found to diverge further either as a sister lineage to (Strict consensus tree), unresolved (Bayesian consensus phylogram) or in a grade (MCC chronogram) with *Simmondsia chinensis* and *Asteropeia micraster*. In all topologies the family Molluginaceae is found to be sister group to the suborder Portulacineae, while the family Anacampserotaceae is found to be the sister-group to the family Cactaceae with *Pereskia* species in basal position in the latter group. All topologies recover a monophyletic subfamily Cactoideae with *Blossfeldia* at the base, but the relationships of the

monophyletic subtribe Trichocereinae are not resolved with high certainty within the tribe Cereeae.

Supplement S10a: Variation of the optimal smoothing parameter selected through the process of PLFB cross-validation; dsD



Supplement S10b: Variation of the optimal smoothing parameter selected through the process of PLBP cross-validation; dsD



Outlook and Future Research

This Ph.D. thesis presents a molecular phylogenetic study of the tribe Cereeae s. l. in order to resolve its generic classification, the pattern of floral evolution and pollination syndromes, and provides age estimation analyses for major clades of cacti. These findings provide a framework for investigating the pattern of diversification of South American cacti in time and space. Contemporary biogeographical studies should be based today on explicit hypotheses that can be tested with appropriate comparative data (e.g., Sanmartin 2012). For instance, a hypothesis concerning the presence of the diverse epiphytic species *Rhipsalis baccifera* (Barthlott 1983) in tropical Africa, Madagascar, and the islands in the Indian Ocean might be formulated as being the result of an old vicariance event associated with the breakup of Gondwana during the Mesozoic. One would find support for this hypothesis when molecular dating information concerning the diversification of this lineage would fall into this period. However, the current age estimation studies (see chapters 3, 4) clearly indicate that the diversification of extant cacti only started, most probably, during the Eocene, and hence, this biogeographical scenario can be firmly rejected.

Concerning the diversification of Cereeae s.l. in space and time we can provide here some first insights. The foundation for future research in this field is clearly now laid down. Cereeae s.l. is mainly distributed in South America and clearly has its origin on this continent. Information on estimated ages for stem lineages and crown groups of the tribe Cereeae s.l. as well as the two prominent subtribes Cereinae and Trichocereinae, and their main distribution areas are presented in this table (age estimates from UCLN analysis based on data set dsA; see chapter 4):

clade	origin (SL) mya	diversification (CG) mya	distribution area
Cereeae	16.2 (11.2-21.5)	12.3 (8.1 - 17.9)	mainly South America, a few taxa in Central America and the Caribbean
Cereinae	12.3 > 10.1	?	mainly NE South America, Caatinga and nearby areas in Brazil
Trichocereinae	10.1 (no HPD)	6.6 (3.8 - 10.3)	mainly Altiplano and foothills of the central Andes of South America

During the late Miocene, a large portion of the lowland parts of the South American continent was flooded by marine transgressions that took place at least twice (i.e., Paranaense sea; Hulka et al. 2006, Marshall et al. 1993). As a result, occasional ruptures in the land connection between East Andes and Eastern Brazil have isolated distinctive lineages on both sides. Lateron, Andean tectonic activities during the Pliocene resulted in a continued increase in elevation (Gregory-Wodzicki 2000, Mosolf et al. 2010). These geomorphological,

as well as climatical events concerning increased seasonality and aridity; Hulka & Heubeck 2010, Mulch et al. 2010), form the biogeographical context in which the diversification of Cereeae s.l. must be explained and understood. Based on phylogenetic analyses and age estimation studies conducted for this Ph.D. thesis (see chapters 1, 4), the split between Cereinae and Trichocereinae can be dated to fall within the period of 12.3 mya (95% HPD up to 17.9; i.e., onset of the diversification of Cereeae s.l.) and 10.1 mya (no HPD available from BEAST analysis). This period of spatial diversification, with Cereinae mainly occurring in North-Eastern Brazil and Trichocereinae being most diverse in the Altiplano and along the foothills of the Andes, might coincide with the formation of the Paranaense sea. At least, temporal information concerning the age of the two lineages are not in disagreement, and hence, such a scenario cannot be rejected. Furthermore, the diversification of the four major subclades of Trichocereinae, containing lineages that either are restricted to the East or to the West of the main Andes is dated to 6.6 mya (95% HPD 3.8 - 10.3 mya), which coincides with the latest phase of Andean orogeny. Along these lines of biogeographical investigations, on the basis of sound phylogenetic analyses and age estimation studies, we can start to better understand the spatial and temporal diversification of Cereeae s.l.

References:

Barthlott, W. 1983. Biogeography and evolution in neo- and palaeotropical Rhipsalinae. In K. Kubitzki [ed.], Dispersal and Distribution, 241–248. Verlag Paul Parey, Hamburg, Germany.

Gregory-Wodzicki, K. M. 2000. Uplift history of the Central and Northern Andes: A review. Geological Society of America Bulletin 112: 1091–1105.

Hulka, C., K.-U., Gräfe, B., Sames, C. E., Uba, C. E. & Heubeck, C. 2006. Depositional setting of the Middle to Late Miocene Yecua Formation of the Chaco Foreland Basin, southern Bolivia. Journal of South American Earth Sciences 21: 135–150.

Hulka, C. & Heubeck, C. 2010. Composition and provenance history of late Cenozoic sediments in southeastern Bolivia: Implications for chaco foreland basin evolution and Andean uplift. Journal of Sedimentary Research 80: 288–299.

Marshall, L. G., Sempere, T. & Gayet, M. 1993. The Petaca (Late Oligocene-Middle Miocene) and Yecua (Late Miocene) formations of the subandean-chaco basin, Bolivia, and their tectonic significance. Documents du Laboratoire de Géologie de Lyon, France 125: 291–301.

Mosolf, J. G., Horton, B. K., Heizler, M. T. & Matos, R. 2010. Unroofing the core of the central Andean fold-thrust belt during focused late Miocene exhumation: evidence from the Tipuani-Mapiri wedge-top basin, Bolivia. Basin Research 23: 346–360.

Mulch, A., Uba, C. E., Strecker, M. R., Schoenberg R. & Page Chamberlain, C. 2010. Late Miocene climate variability and surface elevation in the central Andes. Earth and Planetary Science Letters 290: 173–182.

Sanmartín, I. 2012. Historical Biogeography: Evolution in Time and Space. Evolution: Education and Outreach: 1-14.

Curriculum Vitae

Surname: LENDEL ALEXANDRU

Name: Anita

Date of birth: 11.05.1979

Nationality: Croatian

Education

Secondary school:	I. Gymnasia, Osijek, Croatia, 1997
University:	Josip Juraj Strossmayer University of Osijek, Department of Biology, Croatia, 2002, Master in education of Biology and Chemistry
Master of research thesis:	The genus Arum from the North-Eastern region of Croatia
PhD thesis:	South American cacti in time and space: studies on the diversification of the tribe Cereeae, with particular focus on subtribe Trichocereinae (Cactaceae)

Employed as PhD student at the University of Zurich since May 2006

Awards

Certificate of Excellence, The Poster Award 2005 of the Symposium of the Zürich-Basel Plant Science Center.

Employments during PhD studies:

Assistant in Sukkulenten-Sammlung, Zurich (12.2005 – 03.2006)

Assistant in Herbar, Institute of Systematic Biology, UZH, Zurich (May 2006)

Lab assistant in the research group of Prof. Peter Endress, Institute of Systematic Biology, UZH, Zurich (Dec. 2005 - March 2007 and July 2010 - June 2012)

Professional activities during Ph.D. studies

Teaching assistance:	Biology undergraduate courses BIO113, BIO 229 (2010) and Microtome course (2011), University of Zurich
Seminar organization:	Meet my country Slideshows, departmental seminar series at the Institute of Systematic Botany and Botanic Gardens, University of Zurich (2007-2008)
Public outreach:	Preparing the exposition ” Stammbaum project”, Botanic Garden, University of Zurich (2007) Public lecture within the series of the weekly talks, Botanic Gardens Zurich, Guide at ‘Tree of Life’ exhibition, Zurich Central Station (2009) Fruit sales at the ‘Obstsortenmarkt’, Botanic Gardens Zurich (2007 & 2009) Volunteering for the Lange Nacht der Museen, Botanic Gardens Zurich (2007 & 2009)

Summary

The present Ph.D. thesis investigates the phylogenetic relationships among the major subclades and genera of the tribe Cereeae s.l., conforming to the BCT clade as defined by the first molecular phylogenetic studies of Cactaceae. A focus is placed on the particularly diverse subtribe Trichocereinae. The acquired phylogenies are used to investigate the evolution of several floral characters and pollination syndromes in Cereeae. A prominent part of the present Ph.D. thesis is devoted to the challenges of estimating divergence ages of different lineages of cacti, by investigating challenges related to taxon and fossil calibration sampling.

Chapter 1 provides a phylogenetic hypothesis about the relationships among all relevant representatives of the tribes Cereeae and Trichocereeae using four combined cpDNA markers for maximum likelihood and maximum parsimony analyses. These investigations confirm expanded tribe Cereeae s.l. consists of three distinct clades, here recognized as subtribes Cereinae, Rebutiinae, and Trichocereinae and three orphan lineages recognized as genera *Aylostera* (and possibly *Mediolobivia* as separate genus), *Gymnocalycium*, and *Uebelmannia*. The widely circumscribed genera *Cleistocactus*, *Echinopsis*, *Espostoa* and *Rebutia* are identified as polyphyletic groups that are now split into smaller genera. We use paired-sites tests to investigate whether the molecular data is in conflict to traditional taxonomic concepts and apply generic reclassification conservatively. We propose (1) to recognize *Bolivicereus*, *Borzicactus*, and *Loxanthocereus* in addition to *Cleistocactus*, (2) *Acanthocalycium*, *Lobivia*, *Setiechinopsis*, and *Trichocereus* in addition to *Echinopsis*, (3) *Vatricania* in addition to *Espostoa*, and (4) *Aylostera* and *Weingartia* (possibly also *Mediolobivia* and *Sulcorebuta*) in addition to *Rebutia*. The tribe Cereeae s.l. is basically a South American taxon, with only few species extending ot the Caribbean. Subtribe Cereinae is centered in the arid regions of NE Brazil, while the subtribes Rebutiinae and Trichocerinae are prominent on the Altiplano and the foothills of Argentina, Bolivia, Chile, and Peru, with several East/West Andean transitions.

Chapter 2 investigates evolutionary diversification of several floral characters and pollination syndromes for the tribe Cereeae s.l. Taxonomic studies so far have largely relied on these characters, in addition to growth form differences, but we can easily now illustrate that these features show extensive homoplasy and therefore lead to artificial classification systems. We infer for large parts of the backbone of the Cereeae s.l. phylogeny flowers to be nocturnal, white to pale colored, actinomorphic, appearing most probably in an ordered sequence along the stems, and based on pollination syndrome are most probably visited by bats. We discuss chiropterophilous flowers as possibly representing the ancestral state in the light of it being accessible to all four major guild of pollinators know for cacti (i.e., bees, birds, hawkmoths and bats). Still, we stress the discrepancy between the actual pollination systems and the observed pollination syndromes that form a wide field for further investigations.

Chapter 3 addresses the timing of the origin of cacti and its mode of diversification. Diversification rate shifts among different cactus lineage are investigated and linked closely to global historical events during the Neogene. The two-step approach includes a broadly

sampled phylogeny of angiosperms, being calibrated with multiple fossil constraints, from which, in second step, a series of additional dating analyses are performed with a greatly expanded taxon sampling of cacti. This study found that the cactus lineage diverged from its closest relatives some 35 million years ago, however, major diversification events were found to be more recent, originating some 5-10 Ma. They appear to be contemporary with several other of the world's major succulent plants across multiple continents. This study suggests that trends in the reduction of precipitation and a drop in CO_2 concentration starting from mid-late Miocene have provided strong ecological advantages to further the evolution of succulent syndromes.

Chapter 4 critically evaluates the performance of three sampling approaches as identified in this study (i.e., "distant fossil calibration" approach, "study-group placeholder" approach, and "secondary calibration" approach) for divergence time estimation when the study group is devoid of a reliable fossil record. Our DFC approach, where the molecular data set for the study group is expanded to include relatives with a sufficiently large sample of fossils proves to be superior to a combination of a SGP approach followed by a SC approach. However, the large size of such data sets poses additional analytical challenges. Using UCLN implemented in BEAST we yield for Cactaceae a crown group age of 38.0 Ma (95% HPD 27.4 - 48.7 Ma). Most other estimates fall within the range of the highest posterior density of this most inclusive investigation.

This Ph.D. thesis closes with a brief outlook on the prospects of conducting more detailed biogeographical investigations, which will greatly profit from the well founded phylogenetic investigations as well as from the age estimation studies.

A personal comment: I believe that this Ph.D. thesis provides a sound hypothesis on the circumscription of the historically variable classification of the genera classified Cereeae and Trichocereeae. It shows that morphological traits in cacti traditionally used in classification systems, show extensive homoplasy and should not, therefore, be used for classificatory decisions. It contributes, additionally, a new point of view in explaining the diversity of the currently controversial hypotheses on the origin of Cactaceae by classifying studies they yield from into one of the; or the combination of the three sampling approaches I discuss ("Distant Fossil Calibration approach", "Study Group Placeholder approach" and a "Secondary Calibration approach"). Finally, this thesis demonstrates that the current molecular dating methods might still not be advanced enough to converge in results and to estimate undisputable divergence times in groups of organisms with a scarce fossil record, especially if different sampling approaches and/or taxon sampling densities are applied. More important, it shows that there is a plethora of factors influencing the results in molecular dating analysis and outlines many of the (un)foreseen dangers on the path to successful divergence time estimates. Even for those who disagree with conclusions, I hope that this Ph.D. thesis will provide enough evidence and inspirations for finding, possibly, better solutions.

Zusammenfassung

Die vorgelegte Doktorarbeit untersucht die phylogentetischen Verwandtschaftsbeziehungen der Gattungen und Untergruppen der Tribus Cereeae s.l., in Übereinstimmung mit dem durch vormalige molekular-phylogenetische Studien der Cactaceae definierten BCT *clade*. Das Augenmerk liegt hierbei insbesondere auf der artenreichen Subtribus Trichocereinae. Die im Rahmen der Doktorarbeit erstellten Phylogenien werden dazu benutzt die Evolution verschiedener Blütenmerkmale und von Bestäuber-Syndromen zu untersuchen. Ein Hauptaugenmerk der vorliegenden Doktorarbeit konzentriert sich dabei auf die Frage wann sich einzelne Evolutionslinien innerhalb der Kakteen aufgespaltet haben. Dies wird unter Zuhilfenahme von unterschiedlichen Stichproben von Taxa und Fossilien zur Kalibrierung des Stammbaumes untersucht.

In Kapitel 1 werden die Verwanschaftsbeziehungen verschiedener repräsentativer Arten der Cereeae und Trichocereeae mittels vier cpDNA Markern unter Verwendung von *maximum likelihood* und *maximum parsimony* Analysen untersucht. Diese Untersuchungen haben gezeigt, dass sich die Tribus Cereeae s.l. aus drei verschiedenen *clades*, den Cereinae, Rebutiinae und den Trichocereinae zusammensetzen, sowie die drei nicht höheren Taxa zugeordneten Gattungen *Aylostera* (und möglicherweise *Mediolobivia* als separate Gattung), *Gymnocalycium* sowie *Uebelmannia* umfassen. Die vormalig weit gefassten Gattungen *Cleistocactus*, *Echinopsis*, *Espostoa* und *Rebutia* werden als polyphyletische Gruppen erkannt und werden neu in kleinere Gattungen unterteilt. *Paired-sites test* wurden benutzt um zu untersuchen ob die Ergebnisse der molekularen Untersuchungen in Konflikt zu den traditionellen taxonomischen Konzepten stehen. Die Neuklassifizierung der genannten Gattungen in weitere neue Gattungen wurde dabei möglichst konservativ vorgenommen. Beruhend auf diesen Analysen wird vorgeschlagen: (1) *Bolivicereus*, *Borzicactus*, und *Loxanthocereus* zusätzlich zu *Cleistocactus*, (2) *Acanthocalycium*, *Lobivia*, *Setiechinopsis*, und *Trichocereus* zusätzlich zu *Echinopsis*, (3) *Vatricania* zusätzlich zu *Espostoa,* sowie (4) *Aylostera* und *Weingartia* (möglicherweise auch *Mediolobivia* und *Sulcorebuta*) zusätzlich zu *Rebutia* in eigene Gattungen zu unterteilen. Die Tribus Cereeae s.l. findet sich hauptsächlich in Südamerika, mit nur wenigen Arten welche ihre nördliche Verbreitungsgrenze in der Karibik haben. Die Subtribus Cereinae hat ihr Hauptverbreitungsgebiet in den ariden Gebieten Nordost Brasiliens, während die Rebutiinae und Trichocerinae hauptsächlich im Altiplano, sowie am Fuss der Anden in Argentinien, Bolivien, Chile und Peru vorkommen.

In Kapitel 2 wird die Diversifikation verschiedener Blütenmerkmale und Bestäuber-Syndromen bei den Cereeae s.l. untersucht. Vorrangehende taxonomische Studien hatten sich bisher primär auf diese Merkmale, sowie Unterschiede der Wuchsform, gestützt. Es kann jedoch gezeigt werden, dass genau diese Merkmale stark homoplastisch sind und deren Benützung zu einer künstlichen Klassifikation geführt hat, welche die phylogenetischen Verwandschaftsverhältnisse nicht korrekt wiedergibt. Aus den durchgeführten Untersuchungen schliessen wir ebenfalls, dass die Stammlinie der Cereeae

s.l. nachtaktive, weisse bis leicht farbige, aktinomorphe Blüten aufwies welche vornehmlich von Fledermäusen bestäubt wurden. Diese fledermausbestäubten Blüten als ursprüngliches Merkmal können ebenfalls von allen vier Bestäubergilden besucht zu werden, welche für die Kakteen gängigerweise unterschieden werden (zusätzlich zu Fledermäusen auch Bienen, Vögel, und Nachtschwärmer). Ebenso wird auf den Unterschied zwischen den eigentlichen Bestäubungssystemen und den beobachteten Bestäuber-Syndromen hingewiesen, die ein breites Feld für zukünftige Untersuchungen aufzeigt.

In Kapitel 3 wird das Alter der Entstehung der Kakteen bestimmt. Änderungen der Artbildungsrate verschiedener Abstammungslinien der Kakteen werden hinsichtlich auf globale Ereignisse innerhalb des Neogens untersucht. Ein zweistufiges Verfahren zur Altersschätzung wurde mittels einer gut untersuchten Phylogenie der Blütenpflanzen, kalibriert durch eine Vielzahl von Fossilien, ausgeführt, wo in einem zweiten Schritt mit einer Serie von zusätzlichen Datierungsverfahren die bisher nicht berücksichtigten Taxa der Kakteen mit eingeschlossen wurden. Dabei stellte sich heraus, dass die Kakteen vor ca. 35 Millionen Jahren entstanden sind, die Hauptradiation der Arten jedoch erst vor ca. 5-10 Millionen Jahren stattgefunden hat. Zu diesem Zeitpunkt zeigen auch viele andere sukkulente Pflanzenlinien eine hohe Artbildungsrate. Offensichtlich hat eine Reduktion der Niederschläge und ein Abfall der CO_2 Konzentration in der Atmosphäre ab Mitte bis Ende des Miozäns dazu geführt, dass sukkulente Evolutionslinien vermehrt Arten hervor gebracht haben.

In Kapitel 4 wird die Effizienz der drei von uns unterschiedenen methodischen Ansätze zur Datierung der Entstehungszeit von Evolutionslinien, in Fällen wenn eine Datierung durch einen Mangel an Fossilien erschwert wird, kritisch untersucht. Wir unterscheiden dabei die *distant fossil calibration* Methode, die study-group placeholder Methode, und die *secondary calibration* Methode. Dabei zeigte sich, dass die *distant fossil calibration* Methode, welche eine möglichst grosse Anzahl an Fossilien von weiter entfernt verwandten Arten einschliesst, als bestes Verfahren erweist gegenüber der *study-group placeholder* Methode welche mit der *secondary calibration* Methode kombiniert wird. Die Grösse der dazu benötigten Datensätze führt jedoch zu zusätzlichen analytischen Problemen. Durch Benützung von der UCLN Methode in BEAST konnte das Alter (*crown group age*) der Familie der Cactaceae mit 38 Millionen Jahre bestimmt werden (95% HPD 27.4 - 48.7 Ma). Alle anderen Altersschätzungen zu den Kakteen fallen ebenfalls in diesen Bereich des Vertrauensintervalls.

Diese Doktorarbeit schliesst mit einem kurzen Überblick über die möglichen biogegeographischen Untersuchungen, welche von der vorliegenden phylogenetischen Untersuchung, sowie der Altersbestimmung der Cactaceae profitieren können.